Stanford CS224W: Advanced architectures in Relational Deep Learning

CS224W: Machine Learning with Graphs

Charilaos Kanatsoulis, Stanford University

http://cs224w.stanford.edu

Announcements

- **Colab 3** is due today
- **Homework 3** is due next Thursday 11/13
  - Recitation session recording on Ed
- **Project Milestone** deadline has been pushed back: now due Tuesday 11/11
- **Colab 4** will be released today
- **Homework 2 grades will be released today!**
  - Regrade requests open until 11/13
- **Practice exam will be released today (on Ed)**

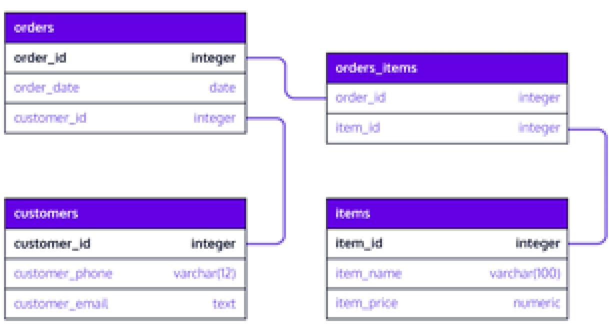
Stanford CS224W: Relational Deep Learning

CS224W: Machine Learning with Graphs

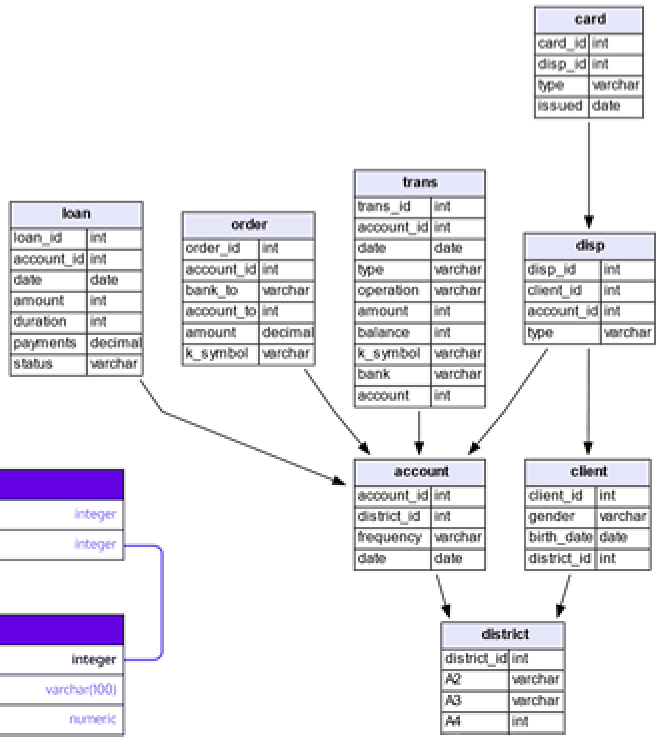
Charilaos Kanatsoulis, Stanford University

http://cs224w.stanford.edu

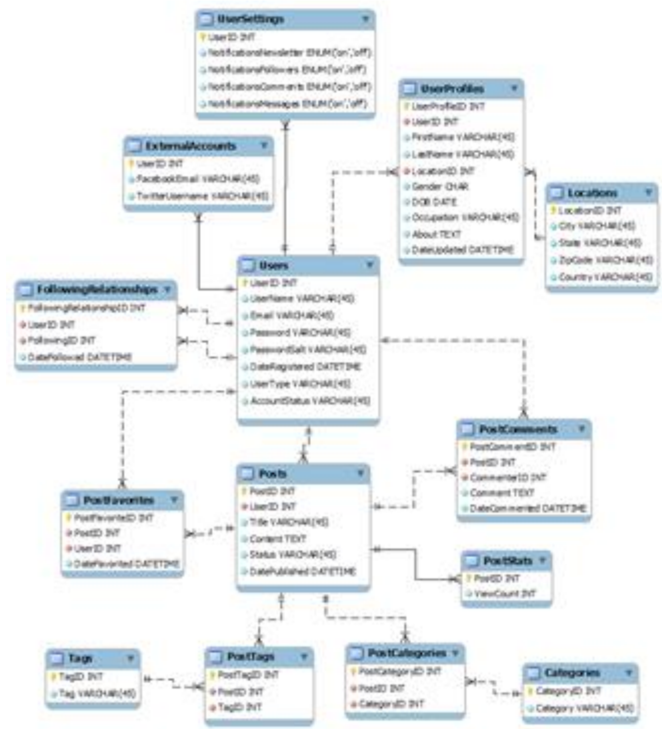
Data stored in Relational Databases



Commerce



Finance



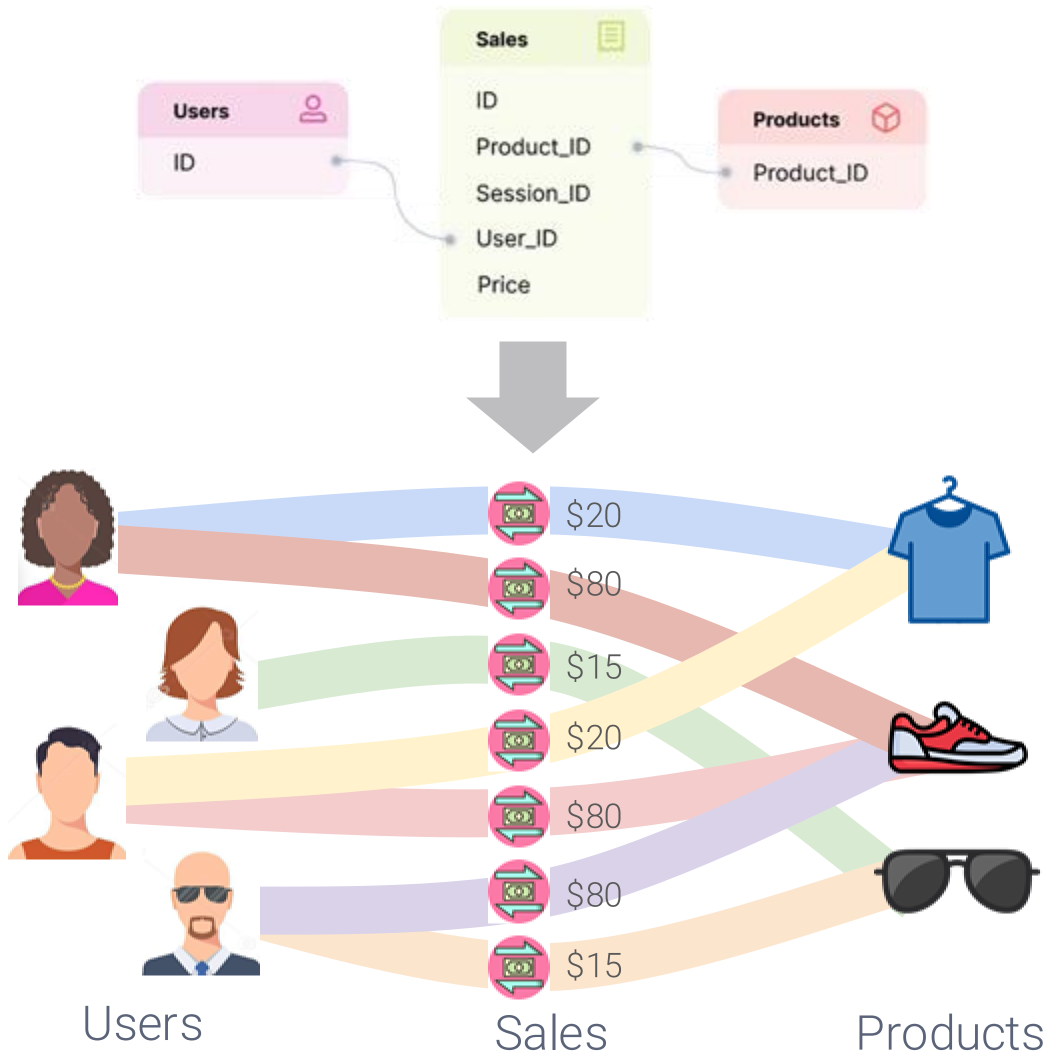
Social Media

Predictions on Relational Data

- Which products will a user purchase in the next 7 days?
- Will an active user churn in the next 90 days?
- What will be the total sales for each product in the next 30 days?

Insight: A Data is a graph!

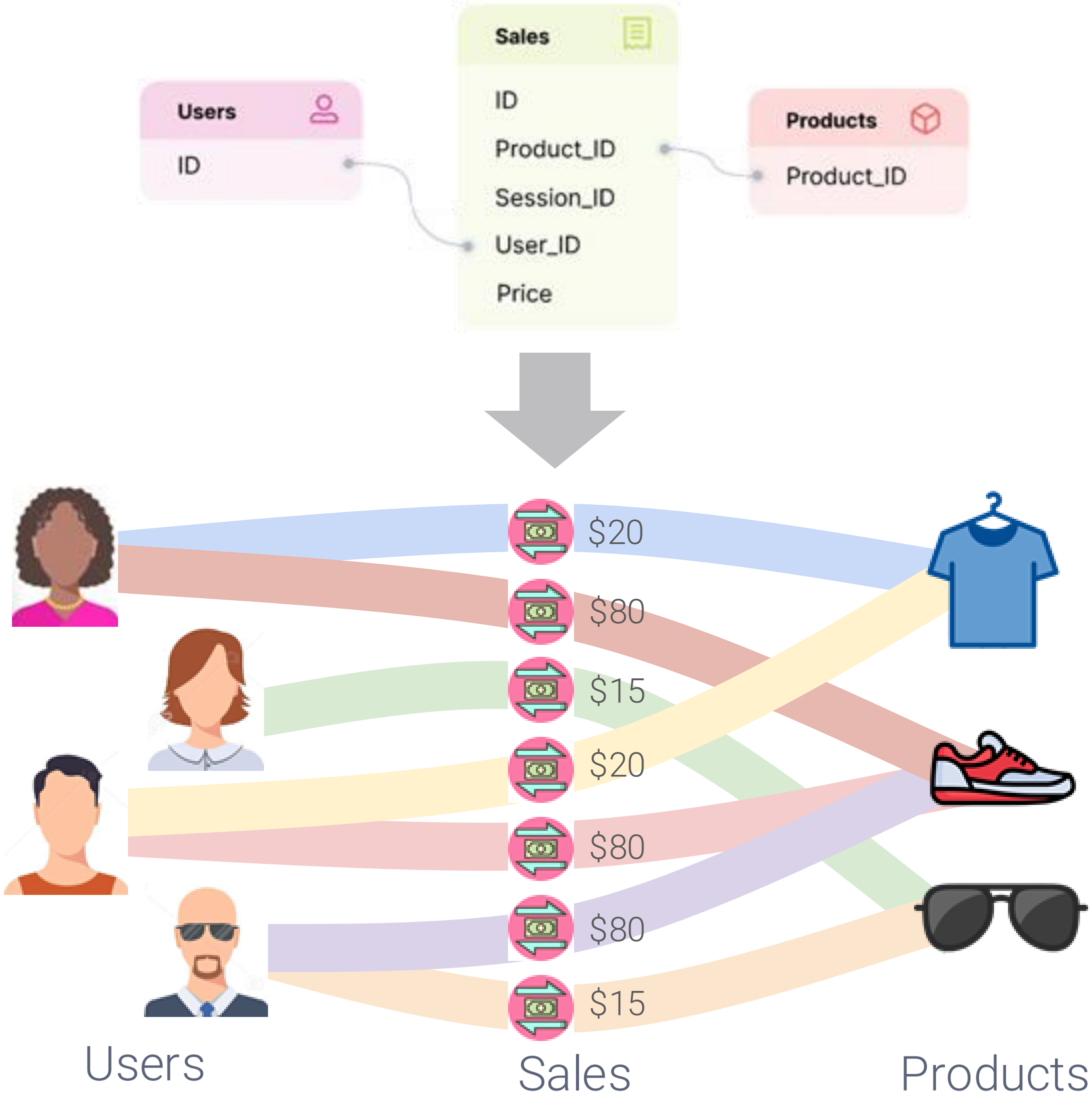
A Database is a graph!



Just do ML on a Graph!

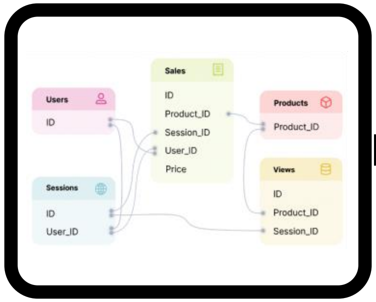
ML in the language of graphs:

- Node-level:
  - Churn
  - Life-time value
  - Next best action
- Link-level:
  - Product affinity
  - Recommendations
- Graph-level:
  - Fraud, money laundering



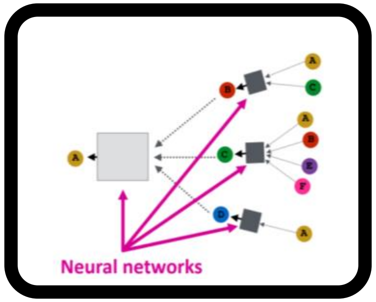
Graph ML Problem Solving Pipeline

Relational DB

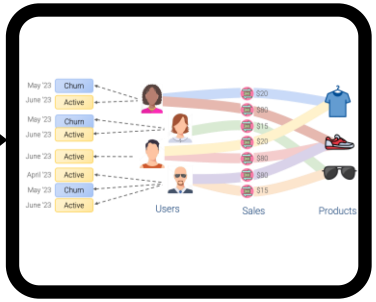


Graph Problem

Graph ML



Solutions



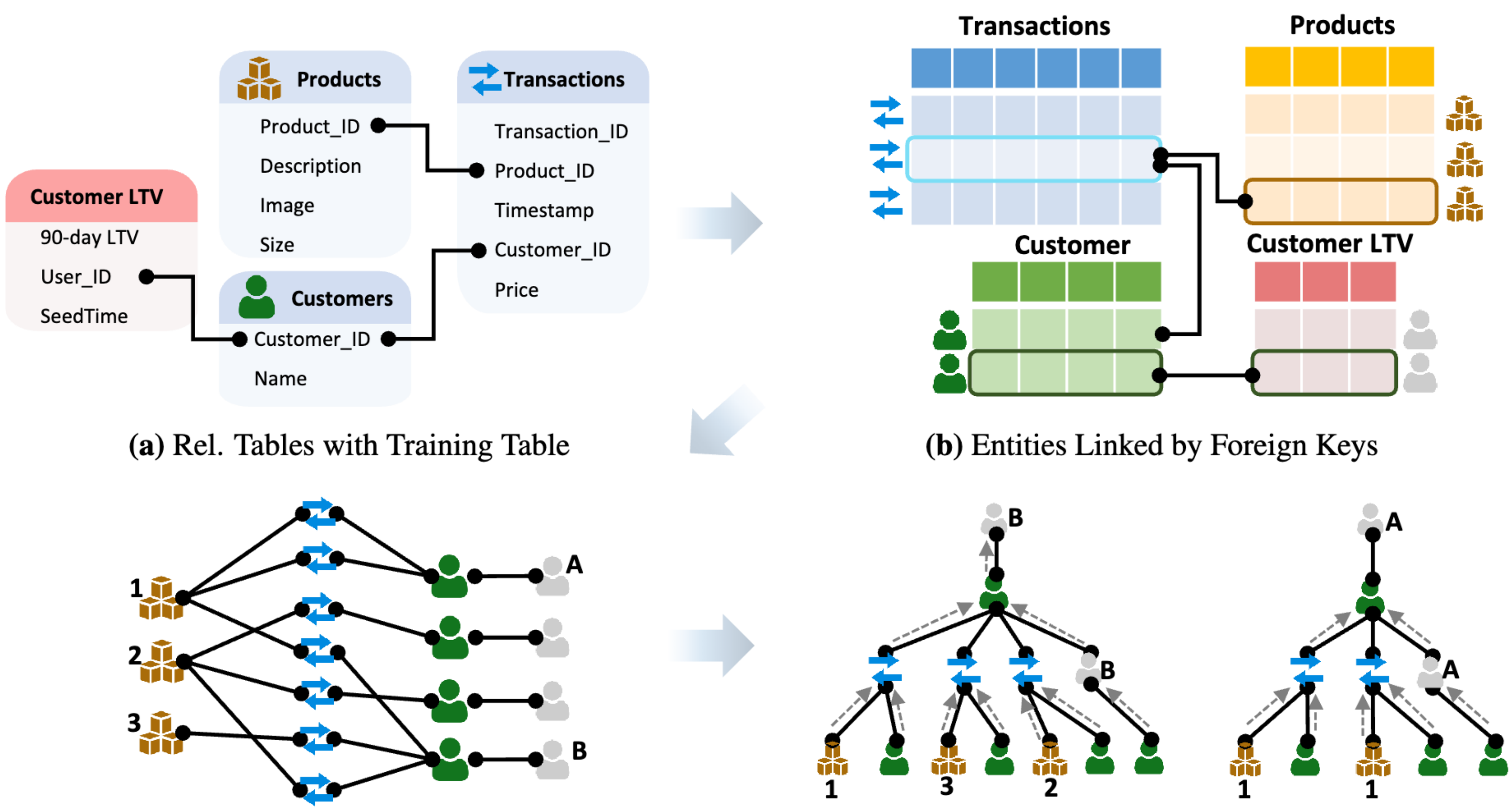
Defining a Task: Training Table

- Training Table: A special table containing training labels
  - (entity ID, time, labels)
    - Classification, Regression, Multi-class

Training Table

Entity ID	Timestamp	Label
99	10172024	1
99	10182024	1
...	...	...
100	10172024	1
100	10182024	0
...	...	...

Relational Deep Learning



(a) Rel. Tables with Training Table

(b) Entities Linked by Foreign Keys

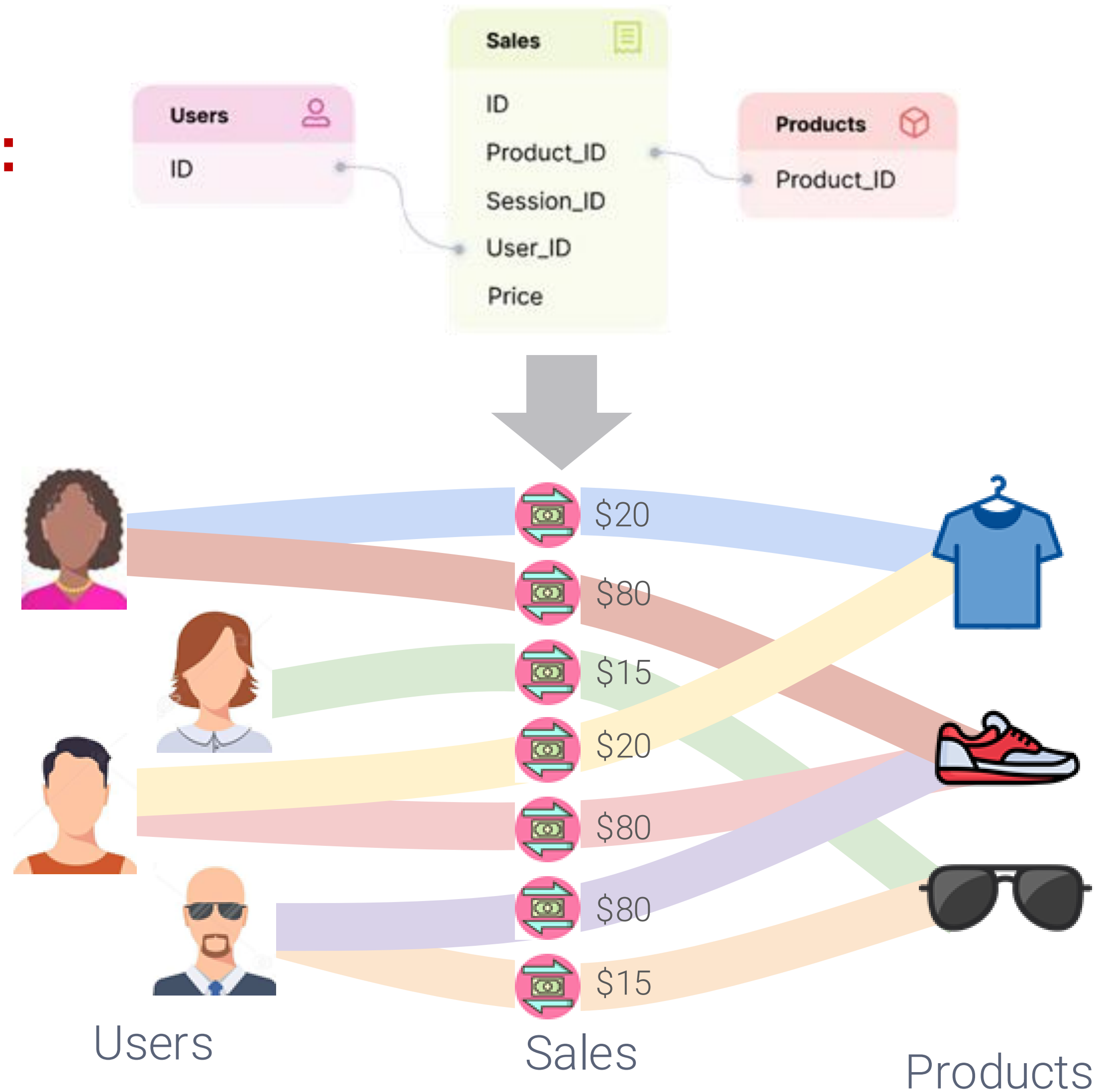
(c) Relational Entity Graph

(d) Graph Neural Network

Relational Entity Graph

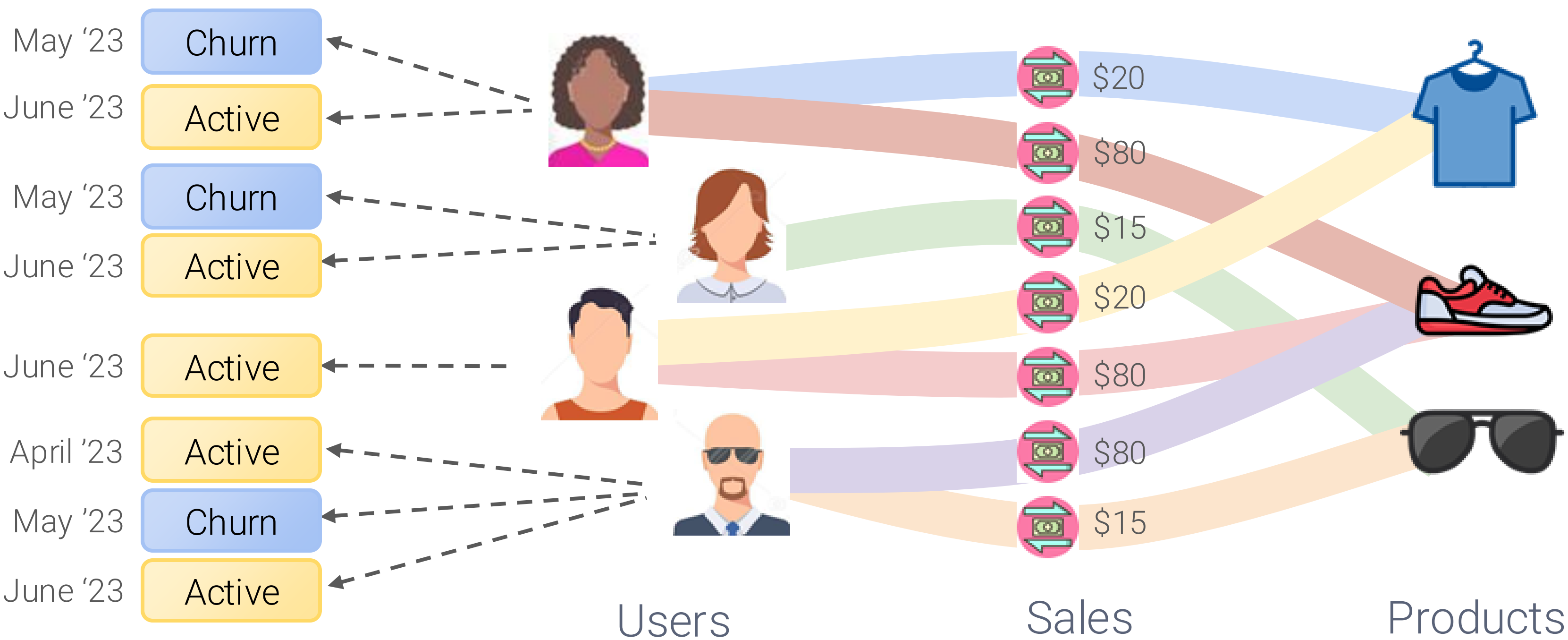
Relational Entity Graph:

Create connections via primary-foreign keys



Connect the Training Table

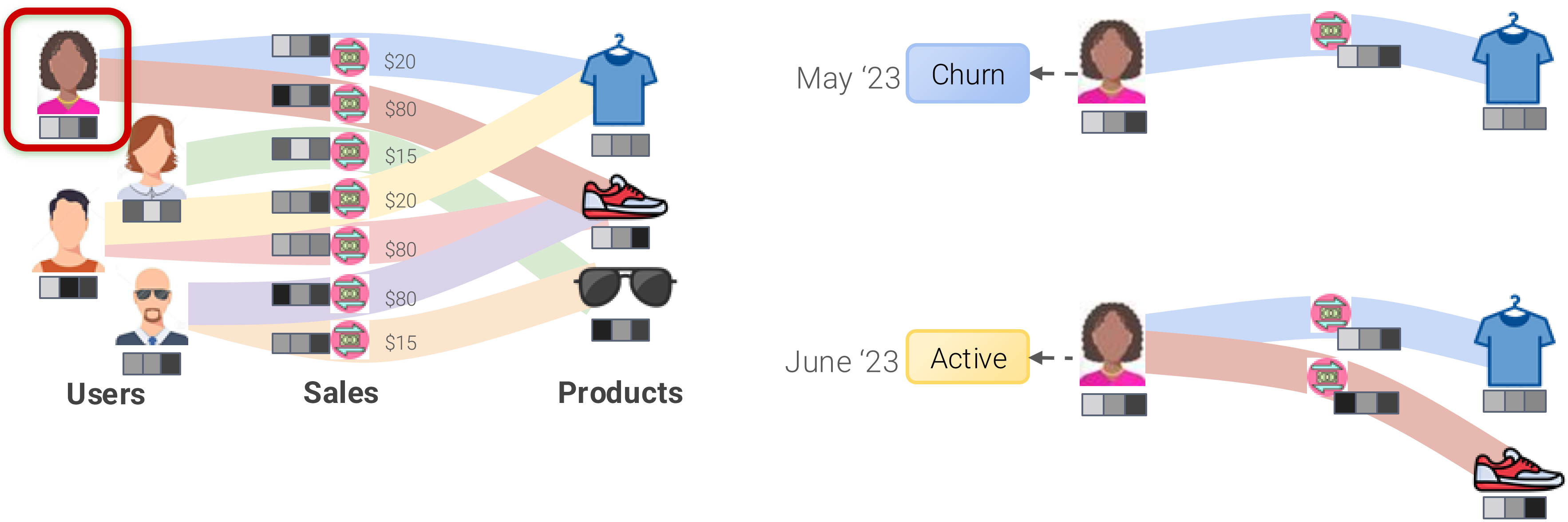
Training labels together with timestamps are attached to the graph



GNN on the Entity Graph

Node's neighborhood defines a computation graph

Nodes learn how to *optimally* use information from neighbors to obtain enhanced node representations



Entity Graph

GNN computation graphs

Why not use an LLM?

Some have tried

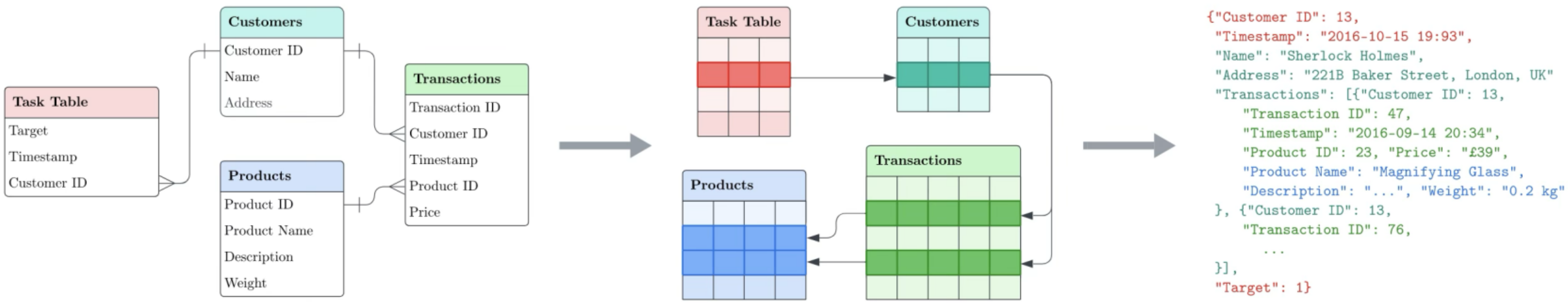


Figure 1: Process of constructing a single example for LLM-based inference.

Tackling prediction tasks in relational databases with LLMs

Marek Wydmuch[*,†] **Łukasz Borchmann**[*] **Filip Graliński**[*,‡]
[*] Snowflake AI Research
{first-name}.{last-name}@snowflake.com
[†] Poznan University of Technology / Poznań, Poland
[‡] Adam Mickiewicz University / Poznań, Poland

Prompt

"Necklace", "product_group_name": "Accessories", "graphical_appearance_no": 1010009, "graphical_appearance_name": "Glittering/Metallic", "colour_group_code": 5, "colour_group_name": "Go
"department_no": 7020, "department_name": "Conscious Exclusive", "index_code": "A", "index_name": "Ladieswear", "index_group_no": 1, "index_group_name": "Ladieswear", "section_no": 82, "
made from recycled metal. Chain pendant with a sparkly stone at the bottom and a tear-shaped pendant covered in spar"}]}, {"t_dat": "2019-11-25T00:00:00", "customer_id": 198761, "article_id": 6
"age": 32.0, "postal_code": "7b4c17acca9c3033b5defe8a1c9eb7b647182d9851dbf92dbf4db3e0e95cb051", "add_transactions": [{"t_dat": "2020-05-04T00:00:00", "customer_id": 198761, "article_
"product_type_name": "Trousers", "product_group_name": "Garment Lower body", "graphical_appearance_no": 1010023, "graphical_appearance_name": "Denim", "colour_group_code": 71, "colo
"perceived_colour_master_name": "Blue", "department_no": 1722, "department_name": "Trouser", "index_code": "A", "index_name": "Ladieswear", "index_group_no": 1, "index_group_name": "La
"5-pocket jeans in washed, superstretch denim with a regular waist, zip fly and button, and skinny legs."}]}, {"t_dat": "2019-11-29T00:00:00", "customer_id": 1348061, "article_id": 68191, "price": 0.
"postal_code": "be5edf195dda1c9409d9c9cc3d32421cc86039819bd20b82bdf6f384660f81d7", "add_transactions": [{"t_dat": "2020-04-11T00:00:00", "customer_id": 1348061, "article_id": 83968,
"product_type_name": "Shirt", "product_group_name": "Garment Upper body", "graphical_appearance_no": 1010001, "graphical_appearance_name": "All over pattern", "colour_group_code": 10, "
"perceived_colour_master_name": "White", "department_no": 7657, "department_name": "Kids Boy Shirt", "index_code": "H", "index_name": "Children Sizes 92-140", "index_group_no": 4, "index_
"Short-sleeved shirt in a cotton weave with a collar, buttons down the front, yoke at the back, and a gently rounded hem."}, {"t_dat": "2020-07-15T00:00:00", "customer_id": 1348061, "article_id": 9
"Blouse", "product_group_name": "Garment Upper body", "graphical_appearance_no": 1010016, "graphical_appearance_name": "Solid", "colour_group_code": 10, "colour_group_name": "White",
"department_no": 1522, "department_name": "Blouse", "index_code": "A", "index_name": "Ladieswear", "index_group_no": 1, "index_group_name": "Ladieswear", "section_no": 15, "section_name
opening at the back with ties at the back of the neck and short, elasticated puff sleeves."}]}, {"t_dat": "2019-12-11T00:00:00", "customer_id": 556304, "article_id": 68191, "price": 0.0338813559322
"6e439ff990d5fc96831fbd32d49825cdc6cb3426c842d7ee0850e109389fd3c9", "add_transactions": [{"t_dat": "2020-03-09T00:00:00", "customer_id": 556304, "article_id": 76457, "price": 0.016932
"product_group_name": "Garment Upper body", "graphical_appearance_no": 1010016, "graphical_appearance_name": "Solid", "colour_group_code": 9, "colour_group_name": "Black", "perceived
1676, "department_name": "Jersey Basic", "index_code": "A", "index_name": "Ladieswear", "index_group_no": 1, "index_group_name": "Ladieswear", "section_no": 16, "section_name": "Womens
a polo neck and long sleeves."}, {"t_dat": "2020-03-09T00:00:00", "customer_id": 556304, "article_id": 76456, "price": 0.0169322033898305, "sales_channel_id": 2, "product_code": 785018, "prod_
1010016, "graphical_appearance_name": "Solid", "colour_group_code": 33, "colour_group_name": "Dark Orange", "perceived_colour_value_id": 2, "perceived_colour_value_name": "Medium Dusty
"A", "index_name": "Ladieswear", "index_group_no": 1, "index_group_name": "Ladieswear", "section_no": 16, "section_name": "Womens Everyday Basics", "garment_group_no": 1002, "garment_g
"2020-03-09T00:00:00", "customer_id": 556304, "article_id": 92998, "price": 0.0254067796610169, "sales_channel_id": 2, "product_code": 856109, "prod_name": "Sweaty", "product_type_no": 27
"Solid", "colour_group_code": 9, "colour_group_name": "Black", "perceived_colour_value_id": 4, "perceived_colour_value_name": "Dark", "perceived_colour_master_id": 5, "perceived_colour_mast
"index_group_name": "Ladieswear", "section_no": 16, "section_name": "Womens Everyday Basics", "garment_group_no": 1002, "garment_group_name": "Jersey Basic", "detail_desc": "Trousers in
"2019-09-25T00:00:00", "customer_id": 179592, "article_id": 68191, "price": 0.0338813559322033, "sales_channel_id": 2, "FN": 1.0, "Active": 1.0, "club_member_status": "ACTIVE", "fashion_news
"2019-12-06T00:00:00", "customer_id": 1067378, "article_id": 68191, "price": 0.0338813559322033, "sales_channel_id": 2, "FN": 1.0, "Active": 1.0, "club_member_status": "ACTIVE", "fashion_new
"2020-01-06T00:00:00", "customer_id": 50910, "article_id": 68191, "price": 0.0338813559322033, "sales_channel_id": 2, "FN": null, "Active": null, "club_member_status": "ACTIVE", "fashion_news_
"2020-07-09T00:00:00", "customer_id": 1063111, "article_id": 68191, "price": 0.0237118644067796, "sales_channel_id": 2, "FN": 1.0, "Active": 1.0, "club_member_status": "ACTIVE", "fashion_new
[{"t_dat": "2019-09-30T00:00:00", "customer_id": 1063111, "article_id": 76458, "price": 0.0169322033898305, "sales_channel_id": 2, "product_code": 785018, "prod_name": "Shenzi LP", "product_
"graphical_appearance_name": "Solid", "colour_group_code": 93, "colour_group_name": "Dark Green", "perceived_colour_value_id": 2, "perceived_colour_value_name": "Medium Dusty", "perceive
"index_name": "Ladieswear", "index_group_no": 1, "index_group_name": "Ladieswear", "section_no": 16, "section_name": "Womens Everyday Basics", "garment_group_no": 1002, "garment_group
"2019-11-29T00:00:00", "customer_id": 418521, "article_id": 68191, "price": 0.0271016949152542, "sales_channel_id": 2, "FN": null, "Active": null, "club_member_status": "ACTIVE", "fashion_news
"2019-10-06T00:00:00", "customer_id": 935373, "article_id": 68191, "price": 0.0338813559322033, "sales_channel_id": 2, "FN": 1.0, "Active": 1.0, "club_member_status": "ACTIVE", "fashion_news
[{"t_dat": "2020-07-14T00:00:00", "customer_id": 935373, "article_id": 99599, "price": 0.0249999999999999, "sales_channel_id": 2, "product_code": 885624, "prod_name": "salt ss", "product_type
"graphical_appearance_name": "Solid", "colour_group_code": 10, "colour_group_name": "White", "perceived_colour_value_id": 3, "perceived_colour_value_name": "Light", "perceived_colour_mast
"Divided", "index_group_no": 2, "index_group_name": "Divided", "section_no": 53, "section_name": "Divided Collection", "garment_group_no": 1005, "garment_group_name": "Jersey Fancy", "deta
{"t_dat": "2020-02-08T00:00:00", "customer_id": 935373, "article_id": 88302, "price": 0.0146779661016949, "sales_channel_id": 1, "product_code": 832361, "prod_name": "Mona", "product_type_
"Solid", "colour_group_code": 12, "colour_group_name": "Light Beige", "perceived_colour_value_id": 1, "perceived_colour_value_name": "Dusty Light", "perceived_colour_master_id": 11, "perceive
"index_group_no": 1, "index_group_name": "Ladieswear", "section_no": 15, "section_name": "Womens Everyday Collection", "garment_group_no": 1005, "garment_group_name": "Jersey Fancy", "
91510, "price": 0.008457627118644, "sales_channel_id": 2, "product_code": 849747, "prod_name": "Fashion offer Reymond 2PCS set", "product_type_no": 270, "product_type_name": "Garment S
"colour_group_name": "Light Yellow", "perceived_colour_value_id": 7, "perceived_colour_value_name": "Medium", "perceived_colour_master_id": 8, "perceived_colour_master_name": "Yellow", "de
"index_group_name": "Baby/Children", "section_no": 41, "section_name": "Baby Boy", "garment_group_no": 1005, "garment_group_name": "Jersey Fancy", "detail_desc": "Set with a vest top and
"customer_id": 935373, "article_id": 40589, "price": 0.0338813559322033, "sales_channel_id": 2, "product_code": 665089, "prod_name": "Carla culotte", "product_type_no": 272, "product_type_na
"colour_group_code": 9, "colour_group_name": "Black", "perceived_colour_value_id": 4, "perceived_colour_value_name": "Dark", "perceived_colour_master_id": 5, "perceived_colour_master_name
"index_group_name": "Ladieswear", "section_no": 15, "section_name": "Womens Everyday Collection", "garment_group_no": 1005, "garment_group_name": "Jersey Fancy", "detail_desc": "Wide, l
935373, "article_id": 85831, "price": 0.0254067796610169, "sales_channel_id": 2, "product_code": 821148, "prod_name": "Corrine", "product_type_no": 258, "product_type_name": "Blouse", "proc
"colour_group_name": "Light Orange", "perceived_colour_value_id": 3, "perceived_colour_value_name": "Light", "perceived_colour_master_id": 3, "perceived_colour_master_name": "Orange", "dep
"section_no": 15, "section_name": "Womens Everyday Collection", "garment_group_no": 1010, "garment_group_name": "Blouses", "detail_desc": "Fitted blouse in a crêpe weave with elastication c
935373, "article_id": 82772, "price": 0.0169322033898305, "sales_channel_id": 2, "product_code": 809874, "prod_name": "Simona", "product_type_no": 254, "product_type_name": "Top", "produc
"colour_group_name": "White", "perceived_colour_value_id": 3, "perceived_colour_value_name": "Light", "perceived_colour_master_id": 9, "perceived_colour_master_name": "White", "department
"section_no": 15, "section_name": "Womens Everyday Collection", "garment_group_no": 1005, "garment_group_name": "Jersey Fancy", "detail_desc": "Shorter, fitted top in ribbed cotton jersey wi
"sales_channel_id": 2, "product_code": 662948, "prod_name": "Osman (1)", "product_type_no": 252, "product_type_name": "Sweater", "product_group_name": "Garment Upper body", "graphical_
"perceived_colour_value_id": 2, "perceived_colour_value_name": "Medium Dusty", "perceived_colour_master_id": 13, "perceived_colour_master_name": "Brown", "department_no": 1626, "departm
"section_name": "Womens Everyday Collection", "garment_group_no": 1003, "garment_group_name": "Knitwear", "detail_desc": "Shorter polo-neck jumper in a rib knit containing some wool with c
"price": 0.0677796610169491, "sales_channel_id": 2, "product_code": 763660, "prod_name": "Premium Poofs down jkt (PQ)", "product_type_no": 262, "product_type_name": "Jacket", "product_g
"colour_group_name": "Greenish Khaki", "perceived_colour_value_id": 4, "perceived_colour_value_name": "Dark", "perceived_colour_master_id": 20, "perceived_colour_master_name": "Khaki gree
"index_group_name": "Baby/Children", "section_no": 41, "section_name": "Baby Boy", "garment_group_no": 1007, "garment_group_name": "Outdoor", "detail_desc": "Jacket in woven fabric with a
"customer_id": 935373, "article_id": 83620, "price": 0.0406610169491525, "sales_channel_id": 2, "product_code": 812683, "prod_name": "Notting Hill", "product_type_no": 265, "product_type_nan
"colour_group_code": 12, "colour_group_name": "Light Beige", "perceived_colour_value_id": 1, "perceived_colour_value_name": "Dusty Light", "perceived_colour_master_id": 11, "perceived_colou
"index_group_name": "Ladieswear", "section_no": 15, "section_name": "Womens Everyday Collection", "garment_group_no": 1013, "garment_group_name": "Dresses Ladies", "detail_desc": "Long
"2019-10-06T00:00:00", "customer_id": 935373, "article_id": 67942, "price": 0.0338813559322033, "sales_channel_id": 2, "product_code": 752814, "prod_name": "Milk RW slack", "product_type_r
"graphical_appearance_name": "Check", "colour_group_code": 9, "colour_group_name": "Black", "perceived_colour_value_id": 4, "perceived_colour_value_name": "Dark", "perceived_colour_mast
"index_group_no": 1, "index_group_name": "Ladieswear", "section_no": 15, "section_name": "Womens Everyday Collection", "garment_group_no": 1009, "garment_group_name": "Trousers", "deta
{"t_dat": "2019-11-28T00:00:00", "customer_id": 140496, "article_id": 68191, "price": 0.0338813559322033, "sales_channel_id": 2, "FN": null, "Active": null, "club_member_status": "ACTIVE", "fash
"2019-12-02T00:00:00", "customer_id": 1062052, "article_id": 68191, "price": 0.0231525423728813, "sales_channel_id": 2, "FN": 1.0, "Active": 1.0, "club_member_status": "ACTIVE", "fashion_new
[{"timestamp": "2020-06-08T00:00:00", "article_id": 68191, "sales": 0.0}, {"timestamp": "2020-05-18T00:00:00", "article_id": 68191, "sales": 0.0}, {"timestamp": "2019-12-16T00:00:00", "article_id":
"sales": 0.0}, {"timestamp": "2020-03-30T00:00:00", "article_id": 68191, "sales": 0.0}, {"timestamp": "2019-10-14T00:00:00", "article_id": 68191, "sales": 0.0}, {"timestamp": "2020-08-03T00:00:00"
68191, "sales": 0.0}, {"timestamp": "2019-12-23T00:00:00", "article_id": 68191, "sales": 0.0}, {"timestamp": "2019-11-11T00:00:00", "article_id": 68191, "sales": 0.0}, {"timestamp": "2020-08-31T00
"2020-03-16T00:00:00", "article_id": 68191, "sales": 0.0237118644067796}, {"timestamp": "2020-07-27T00:00:00", "article_id": 68191, "sales": 0.0}, {"timestamp": "2020-07-06T00:00:00", "article_
"article_id": 68191, "sales": 0.0}, {"timestamp": "2020-06-01T00:00:00", "article_id": 68191, "sales": 0.0237118644067796}, {"timestamp": "2020-06-22T00:00:00", "article_id": 68191, "sales": 0.0},
"2020-02-17T00:00:00", "article_id": 68191, "sales": 0.0}, {"timestamp": "2020-02-03T00:00:00", "article_id": 68191, "sales": 0.0}, {"timestamp": "2020-02-10T00:00:00", "article_id": 68191, "sales"
0.0338813559322033}, {"timestamp": "2019-11-25T00:00:00", "article_id": 68191, "sales": 0.17918644067796569}, {"timestamp": "2020-03-02T00:00:00", "article_id": 68191, "sales": 0.0}, {"timest
{"timestamp": "2020-08-10T00:00:00", "article_id": 68191, "sales": 0.0}, {"timestamp": "2019-11-04T00:00:00", "article_id": 68191, "sales": 0.0338813559322033}, {"timestamp": "2020-08-17T00:0
"article_id": 68191, "sales": 0.0}, {"timestamp": "2019-10-28T00:00:00", "article_id": 68191, "sales": 0.0}, {"timestamp": "2019-10-21T00:00:00", "article_id": 68191, "sales": 0.0}, {"timestamp": "202
"2019-09-30T00:00:00", "article_id": 68191, "sales": 0.0338813559322033}, {"timestamp": "2019-11-18T00:00:00", "article_id": 68191, "sales": 0.0322881355932203}, {"timestamp": "2019-09-09T
"article_id": 68191, "sales": 0.0}, {"timestamp": "2020-01-20T00:00:00", "article_id": 68191, "sales": 0.0}, {"timestamp": "2020-04-27T00:00:00", "article_id": 68191, "sales": 0.0}, {"timestamp": "202
"2019-12-09T00:00:00", "article_id": 68191, "sales": 0.0338813559322033}, {"timestamp": "2019-10-07T00:00:00", "article_id": 68191, "sales": 0.0}, {"timestamp": "2020-01-06T00:00:00", "article_

Why not LLMs?

- LLMs are not trained to understand and effectively learn from the relational nature of Databases
- Database prediction is **NOT** sequence modeling
- LLMs are **NOT** trained to predict the future
- Context size is **limited**

RDL gains over LLM

Entity classification results (ROC-AUC(%), higher is better) on RELBENCH test set. Best values are in bold.

Dataset	Task	ICL Llama 3.2 3B	RELGNN(ours)	Relative Gain
rel-amazon	user-churn	62.55	**70.99**	13%
	item-churn	73.41	**82.64**	13%
rel-avito	user-visits	53.36	**66.18**	24%
	user-clicks	54.07	**68.23**	26%
rel-event	user-repeat	70.11	**79.61**	14%
	user-ignore	68.65	**86.18**	26%
rel-f1	driver-dnf	**80.03**	75.29	−6%
	driver-top3	**87.11**	85.69	−2%
rel-hm	user-churn	63.81	**70.93**	11%
rel-stack	user-engagement	81.23	**90.75**	12%
	user-badge	79.99	**88.98**	11%
rel-trial	study-outcome	59.17	**71.24**	20%

- Stay tuned for more discussion on LLMs later!

Stanford CS224W: Next-Generation Architectures for RDL

CS224W: Machine Learning with Graphs
Charilaos Kanatsoulis, Stanford University
http://cs224w.stanford.edu

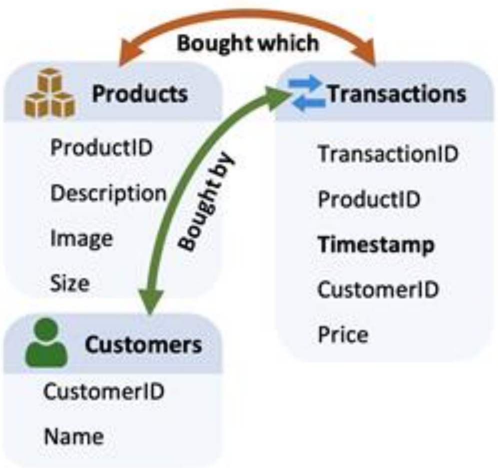
The need for specialized models

Relational Databases: **multi-modal**, **multi-scale** information

The need for specialized models

Relational Databases: **multi-modal**, **multi-scale** information

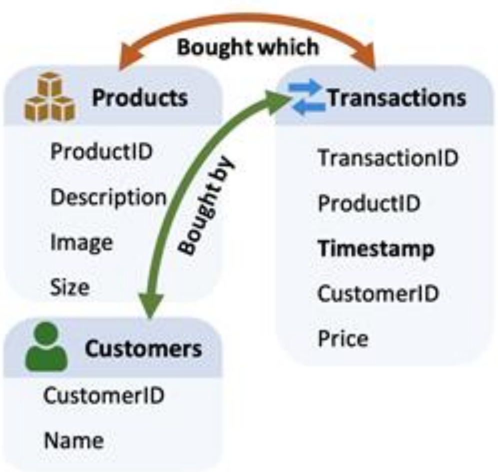
Entity level: Text, image, numerical, categorical, timestamp



The need for specialized models

Relational Databases: **multi-modal**, **multi-scale** information

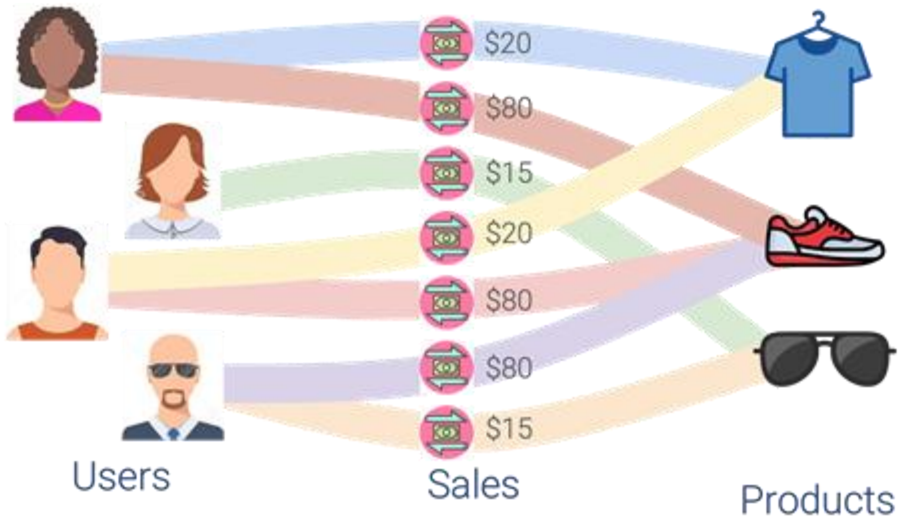
Entity level: Text, image, numerical, categorical, timestamp



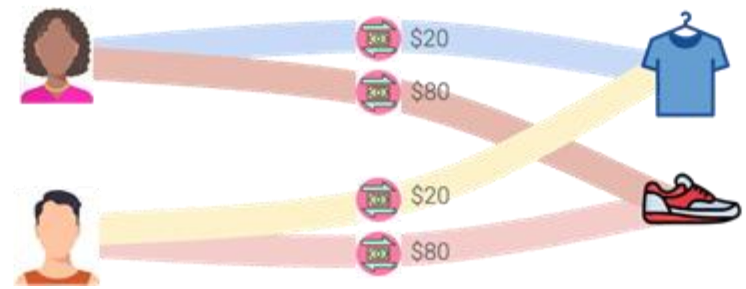
Database level: Relational structure, Temporal Structure

Relational DBs structure

Tripartite patterns



Cycle patterns

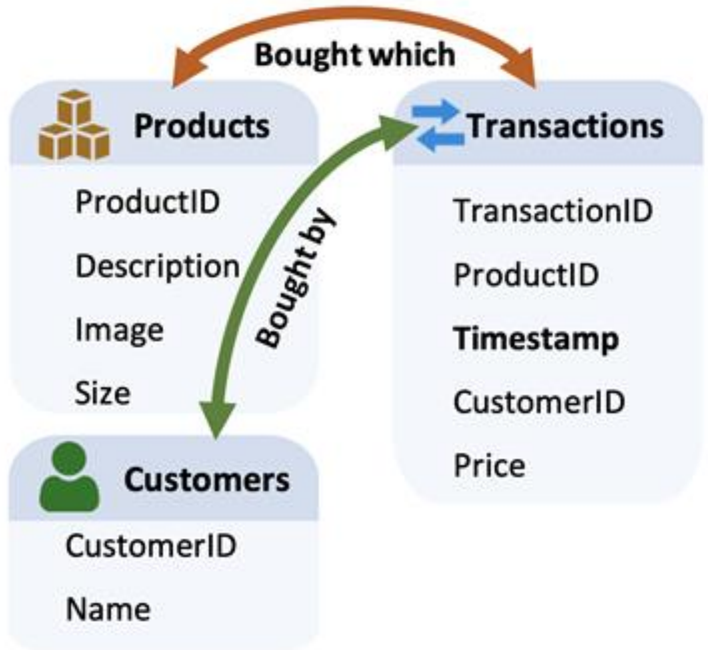


Structure changes over time!

Stanford CS224W: RelGNN: Composite Message-Passing for RDL

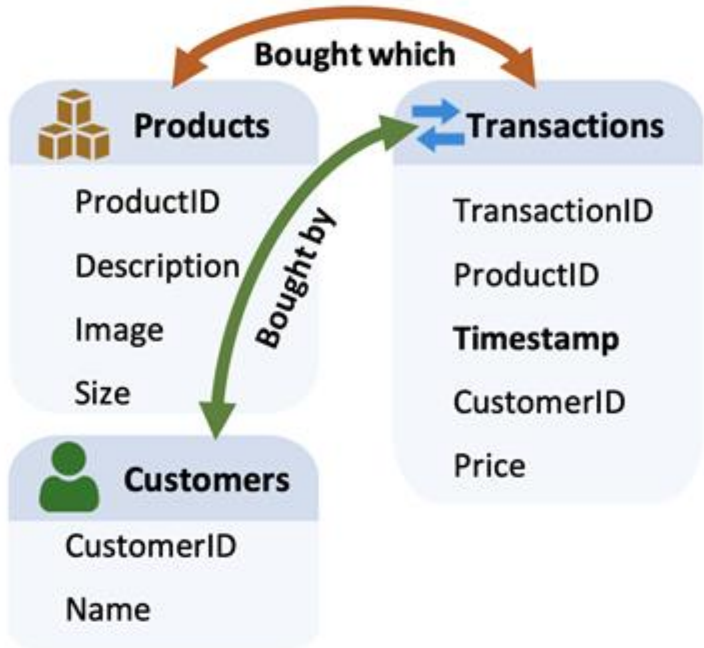
CS224W: Machine Learning with Graphs
Charilaos Kanatsoulis, Stanford University
http://cs224w.stanford.edu

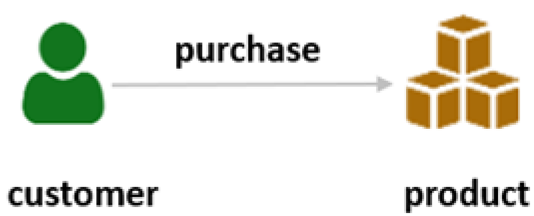
Relational DBs structure



standard heterogeneous graph

Relational DBs structure





standard heterogeneous graph

relational entity graph

Observation

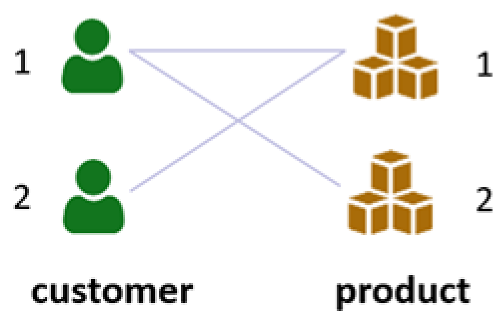
- **Edges in relational entity graphs are defined by primary–foreign key links that merely record table connectivity without semantics.**

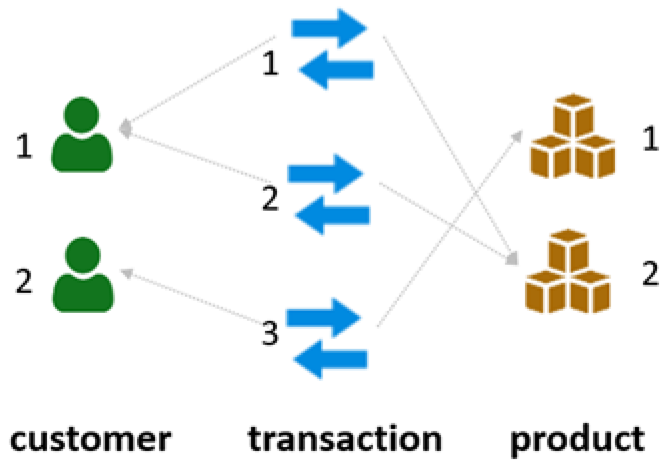
standard heterogeneous graph

relational entity graph

Observation

- Edges in relational entity graphs are **defined by primary–foreign key links** that merely record table connectivity **without semantics**.
- **Junction tables** are introduced, decomposing each many-to-many association into **a pair of one-to-many links**.





RELBENCH

RELATIONAL DEEP LEARNING BENCHMARK

Stanford CS224W: RELBENCH

CS224W: Machine Learning with Graphs

Jure Leskovec, Stanford University

http://cs224w.stanford.edu

RelBench Datasets

7 Diverse Datasets

E-Commerce

- rel-amazon
- rel-avito
- rel-hm

Social

- rel-event
- rel-stack

Sports

- rel-f1

Medical

- rel-trial

Junction Tables in Schema Graph

rel-amazon

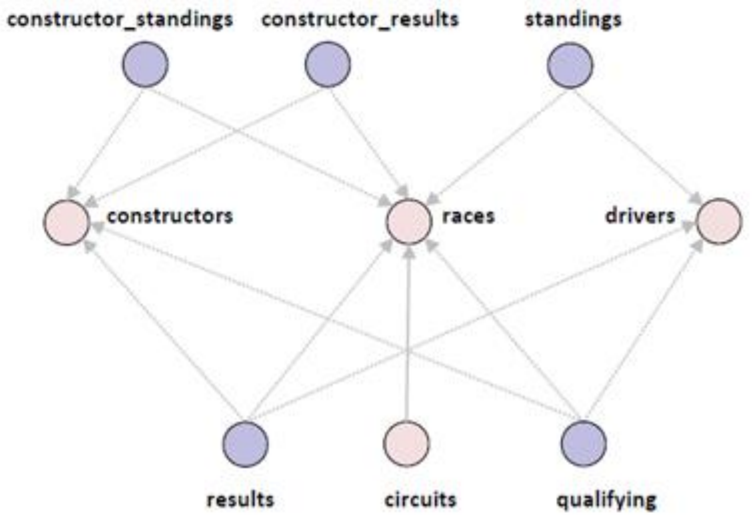
rel-hm

Junction Tables in Schema Graph

rel-amazon

rel-hm

rel-F1

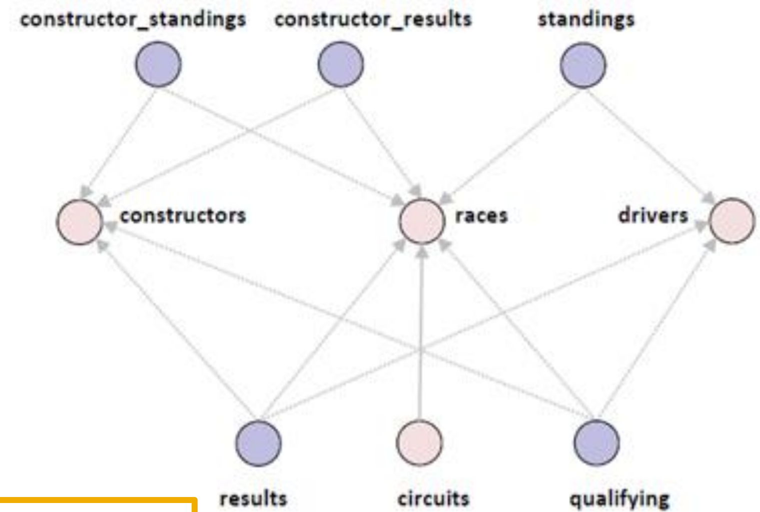


Junction Tables in Schema Graph

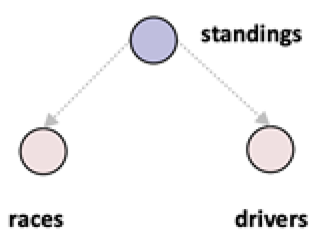
rel-amazon

rel-hm

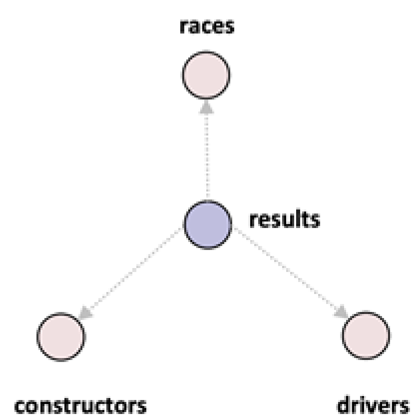
rel-F1



bridge structure

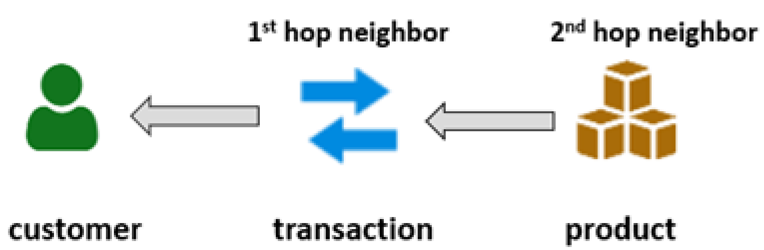


hub structure

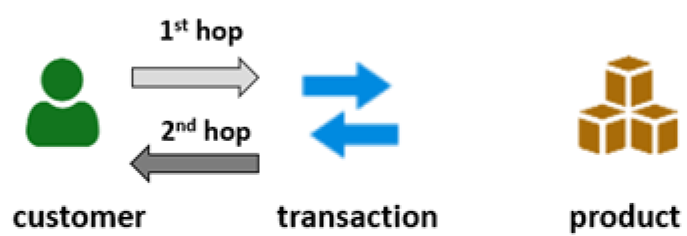


Limitation

- (1) ***Imbalance***: **junction node are 1-hop neighbors and more informative node become 2-hop.**
- (2) ***Redundancy***: **information from the source node is passed to the junction node in the first hop and routed back to itself in the second, aggregated with the junction and 2-hop node information.**



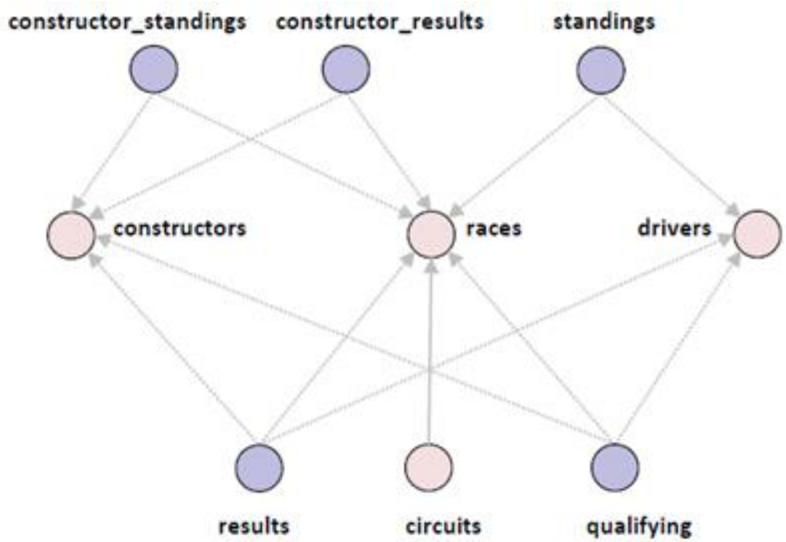
Imbalance



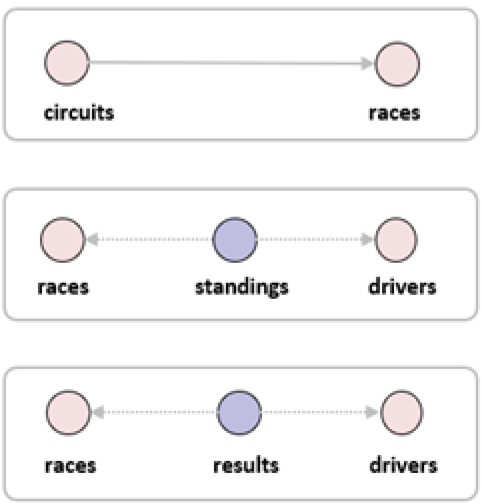
Redundancy

RelGNN: Atomic Routes

- **Informal: An atomic route is a simple path between pink node-types.**
- **Atomic routes can be derived automatically without manual intervention or domain-expert knowledge.**



Primary-foreign Key Relation of rel-f1 Dataset



Example of Atomic Routes

Tianlang Chen, **Charilaos Kanatsoulis**, and Jure Leskovec. "RelGNN: Composite Message Passing for Relational Deep Learning", *In Forty-second International Conference on Machine Learning (ICML), 2025*

RelGNN: Composite Message Passing

- **Composite Message Passing**: aggregates messages along atomic routes in a single step.

$$\mathbf{m}^{(l+1)}_{(\mathtt{dst},\mathtt{mid},\mathtt{src})} = \mathrm{AGGR}(\mathbf{h}^{(l)}_{\mathtt{dst}}, \{\{\mathrm{FUSE}(\mathbf{h}^{(l)}_{\mathtt{mid}}, \mathbf{h}^{(l)}_{\mathtt{src}})\}\})$$

- **Instantiation example:**

$$\mathrm{FUSE}(\mathbf{h}^{(l)}_{\mathtt{mid}}, \mathbf{h}^{(l)}_{\mathtt{src}}) = \mathbf{W}_1\mathbf{h}^{(l)}_{\mathtt{mid}} + \mathbf{W}_2\mathbf{h}^{(l)}_{\mathtt{src}}.$$

$$\mathrm{AGGR}(\mathbf{h}^{(l)}_{\mathtt{dst}}, \{\{\mathbf{h}^{(l)}_{\mathtt{fuse}}\}\}) = \mathbf{W}_{\mathrm{proj}}\mathbf{h}^{(l)}_{\mathtt{dst}} + \sum_{\mathtt{fuse}\in\mathcal{N}(\mathtt{dst})} \alpha_{\mathtt{dst},\mathtt{fuse}}\mathbf{W}_V\mathbf{h}^{(l)}_{\mathtt{fuse}}$$

Standard Message Passing

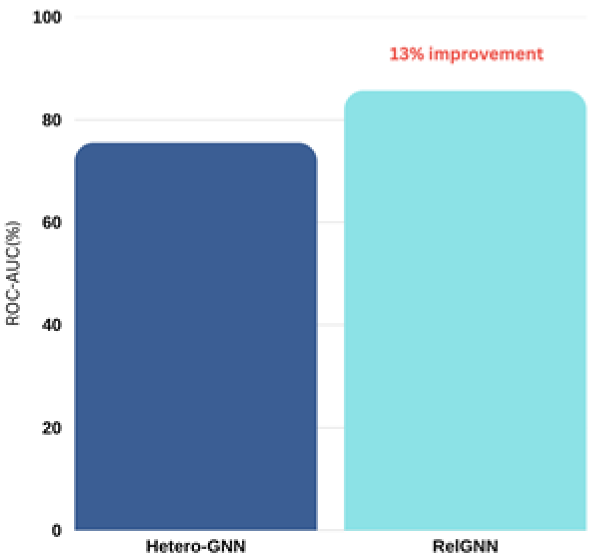
Composite Message Passing

Tianlang Chen, **Charilaos Kanatsoulis**, and Jure Leskovec. "RelGNN: Composite Message Passing for Relational Deep Learning", *In Forty-second International Conference on Machine Learning (ICML), 2025*

RelGNN results

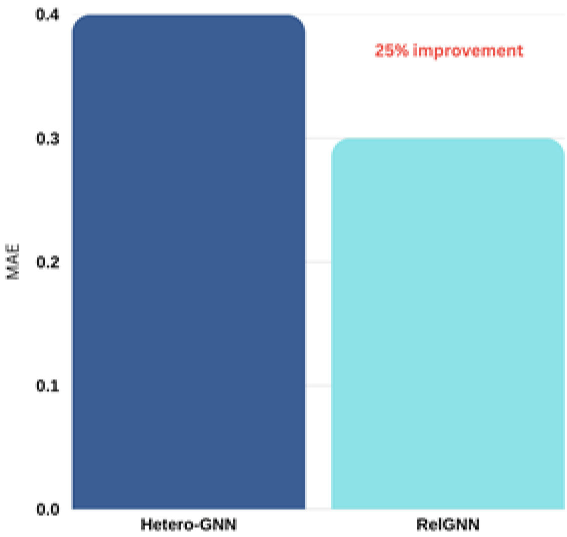
rel-F1

Predict if the driver will qualify in the top-3 for a race in the next 1 month.



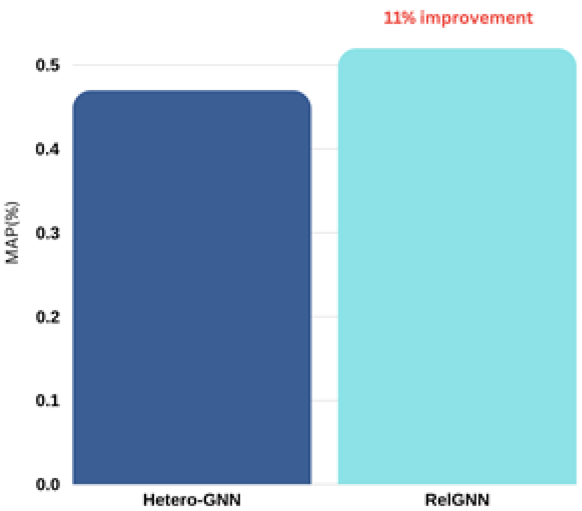
rel-trial

Predict the success rate of a trial site in the next 1 year.



rel-amazon

Predict the list of distinct items each customer will purchase and give a detailed review in the next 3 months.



- **RelGNN achieves SOTA performance on the vast majority of tasks from RelBench, with improvements of up to 25%.**

Tianlang Chen, **Charilaos Kanatsoulis**, and Jure Leskovec. "RelGNN: Composite Message Passing for Relational Deep Learning", *In Forty-second International Conference on Machine Learning (ICML), 2025*

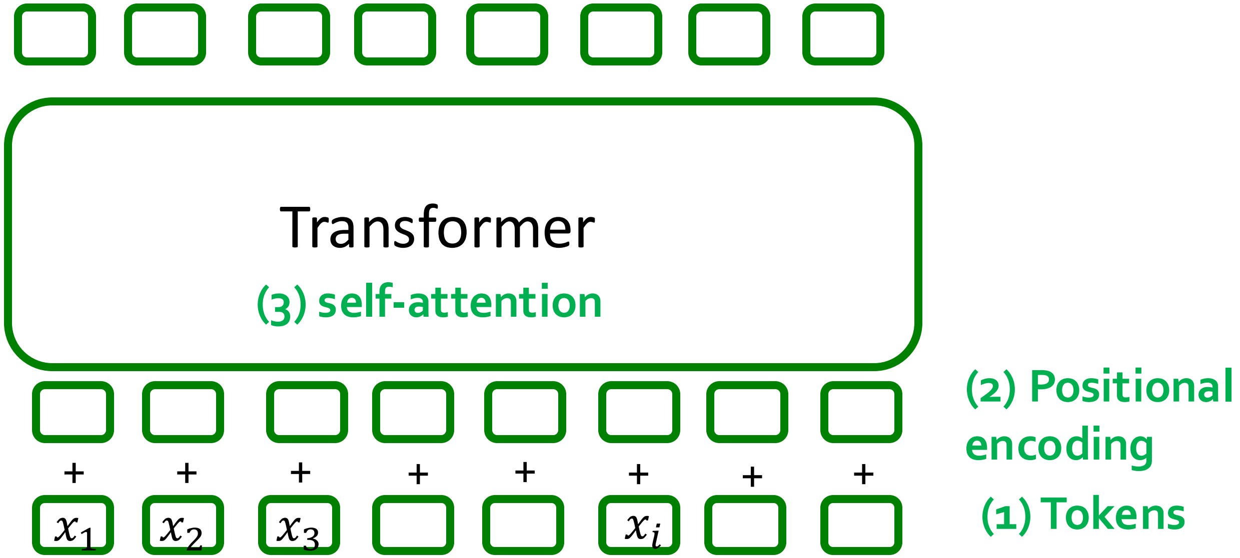
Stanford CS224W: RelGT: Relational Graph Transformer

CS224W: Machine Learning with Graphs
Jure Leskovec, Stanford University
http://cs224w.stanford.edu

Components of a Transformer

- Key components of Transformer
  - **(1) tokenizing**
  - **(2) positional encoding**
  - **(3) self-attention**

- **Key question:** What should these be for a graph input?

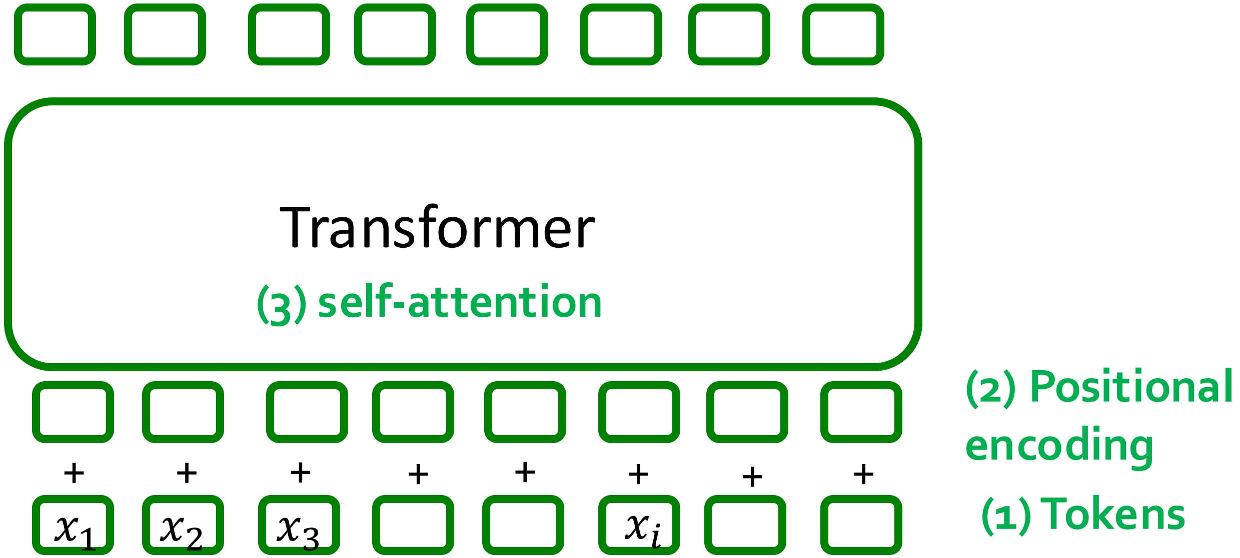


Components of a Transformer

- Key components of Transformer
  - **(1) tokenizing**
  - **(2) positional encoding**
  - **(3) self-attention**

How to chose these for graph data?

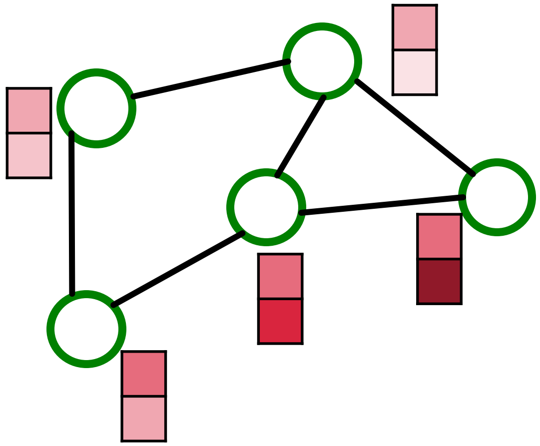
- **Key question:** What should these be for a graph input?



Processing Graphs with Transformers

- A graph Transformer must take the following inputs:
  - **(1) Node features?**
  - **(2) Adjacency information?**
  - **(3) Edge features?**

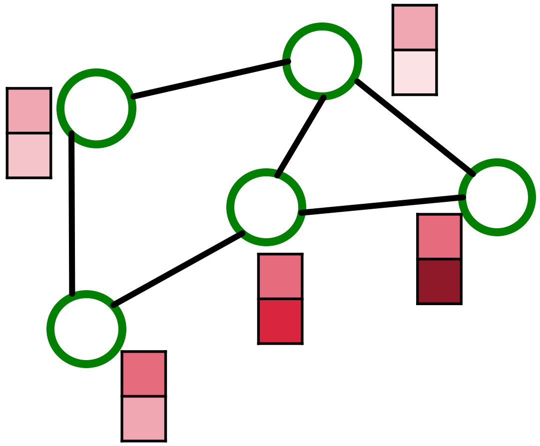
- Key components of Transformer
  - **(1) tokenizing**
  - **(2) positional encoding**
  - **(3) self-attention**



Processing Graphs with Transformers

- A graph Transformer must take the following inputs:
  - (1) Node features?
  - (2) Adjacency information?
  - (3) Edge features?

- Key components of Transformer
  - (1) tokenizing
  - (2) positional encoding
  - (3) self-attention

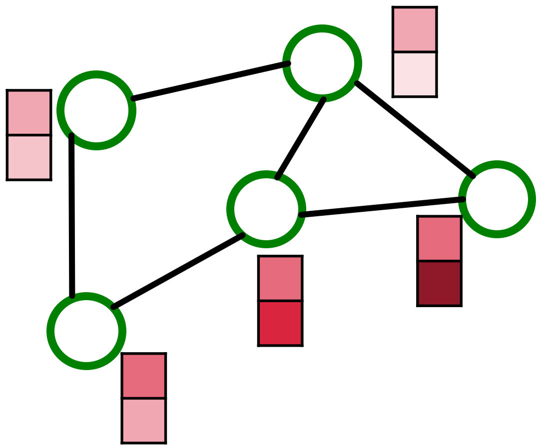


- There are many ways to do this
- Different approaches correspond to different "matchings" between graph inputs **(1), (2), (3)** transformer components **(1), (2), (3)**

Processing Graphs with Transformers

- A graph Transformer must take the following inputs:
  - **(1) Node features?**
  - **(2) Adjacency information?**
  - **(3) Edge features?**

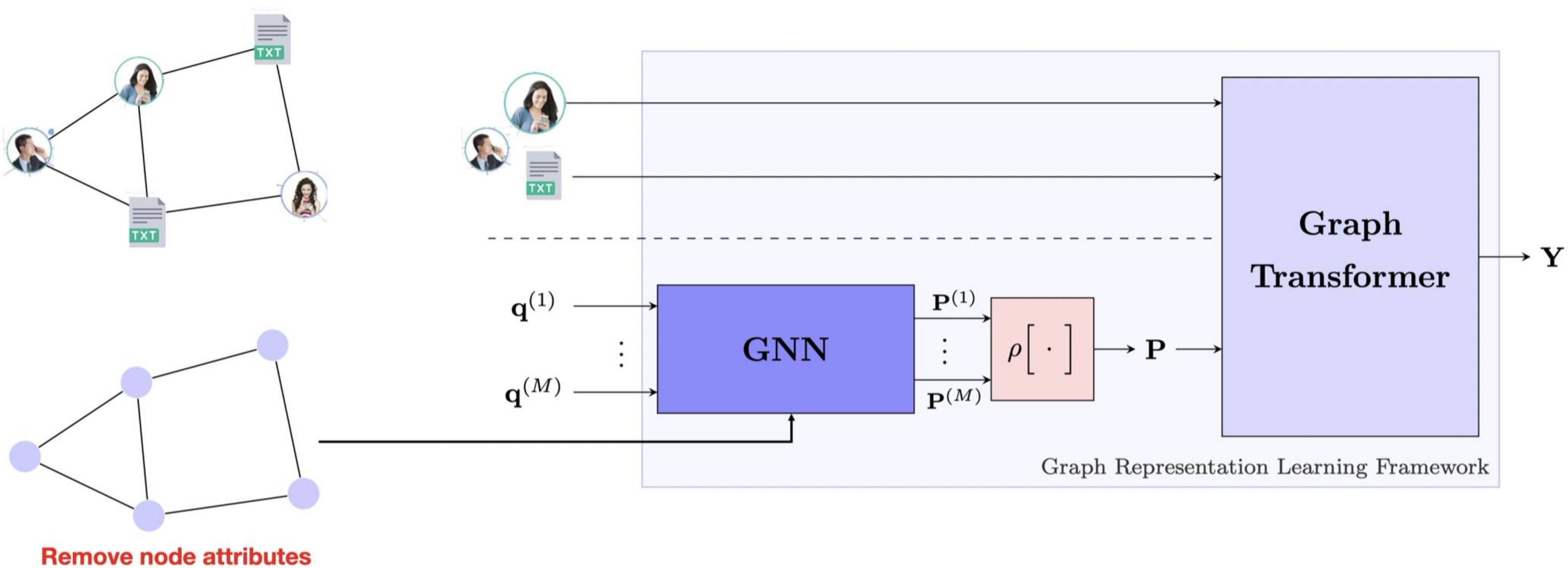
- Key components of Transformer
  - **(1) tokenizing**
  - **(2) positional encoding**
  - **(3) self-attention**



- There are many ways to do this
- Different approaches correspond to different "matchings" between graph inputs **(1), (2), (3)** transformer components **(1), (2), (3)**

PEARL + Transformer

- **How to use PEARL: in practice?**
  - Step 1: Sample node ids from a probability distribution.
  - Step 2: Process each set of node samples independently via a GNN.
  - Step 3: Summarize the outputs via empirical expectation.
  - Step 4: concatenate PEARL embeddings with node features X.
  - Step 5: pass through main GNN/Transformer as usual.
  - Step 6: Backpropagate gradients to train PEARL + Prediction model jointly.

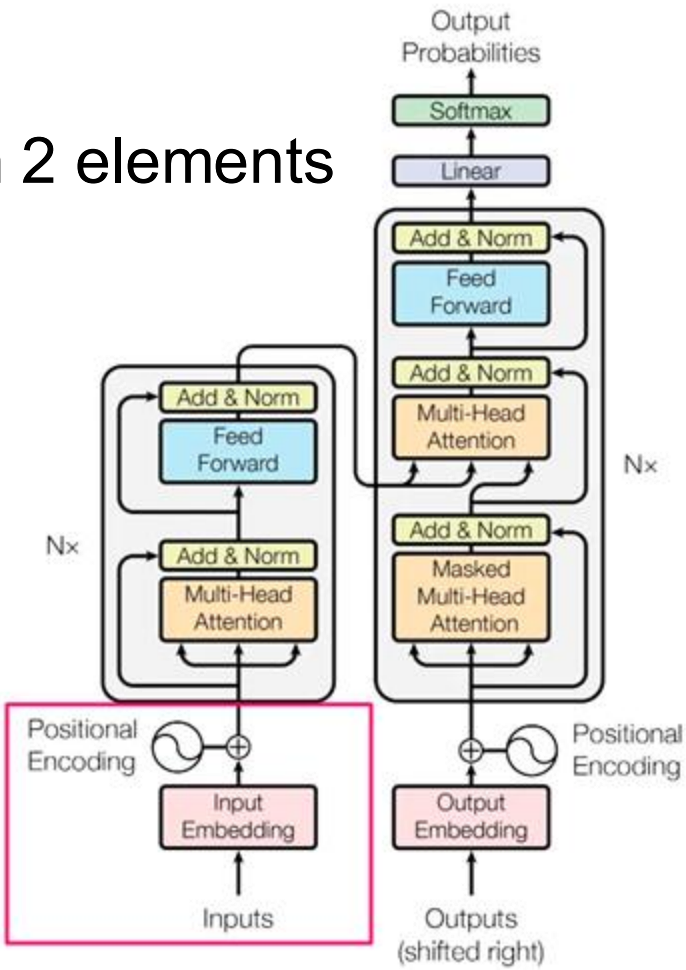


Observation

(Text) Transformers

- An input token is represented with 2 elements
  - Token id, position

    feature structure

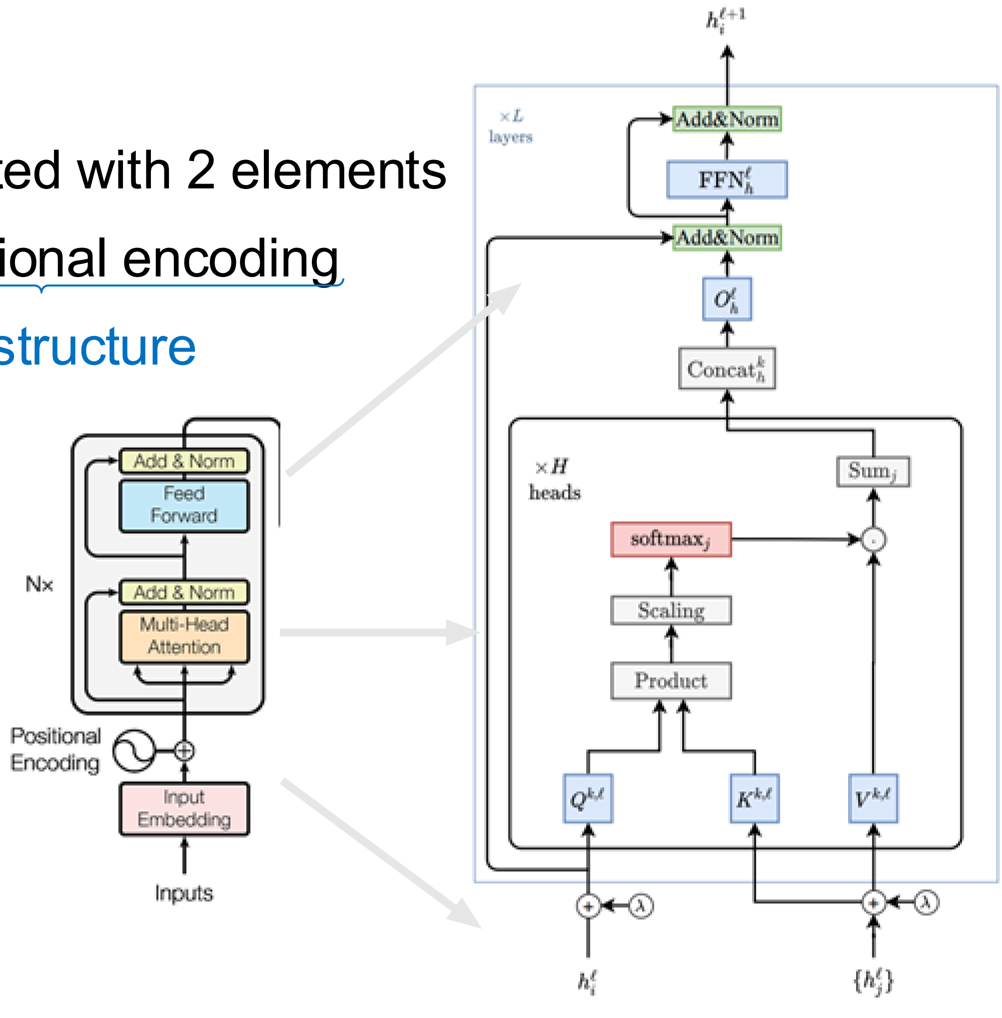


Observation

Graph Transformers

- An input token is represented with 2 elements
  - ○ Node feature, graph positional encoding

feature structure



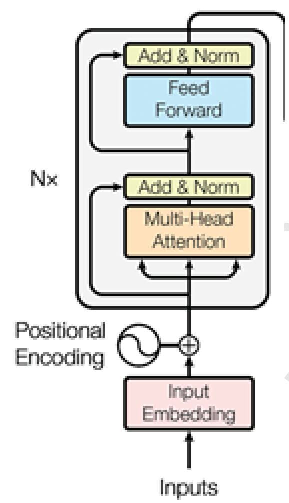
Graph Transformer Layer

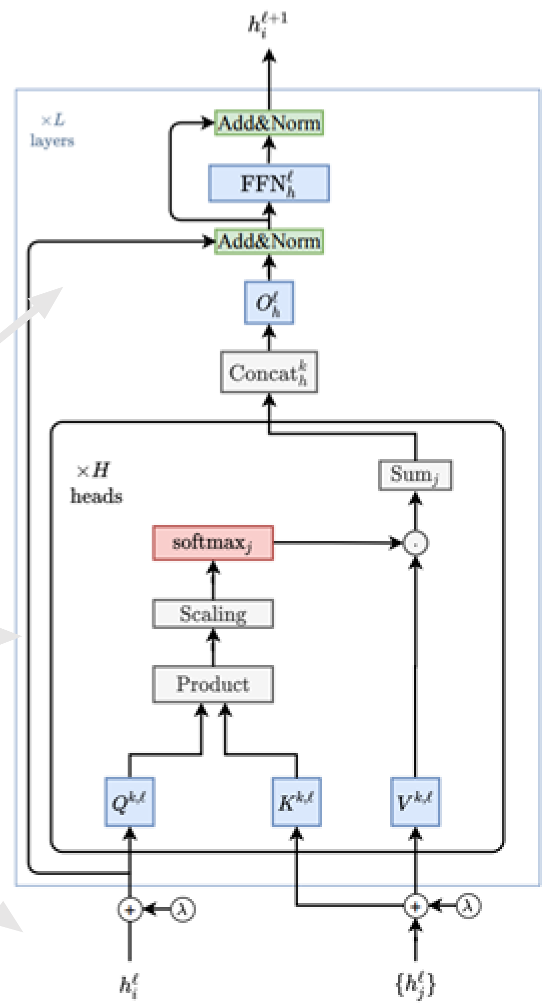
Observation

Graph Transformers

- An input token is represented with 2 elements
  - $\underbrace{\text{Node feature}}_{\text{feature}}$, $\underbrace{\text{graph positional encoding}}_{\text{structure}}$

- Most GTs are designed for homogeneous, static graphs
- However, for additional RDL complexity such as **heterogeneity** and **temporality**, this design needs to be improved





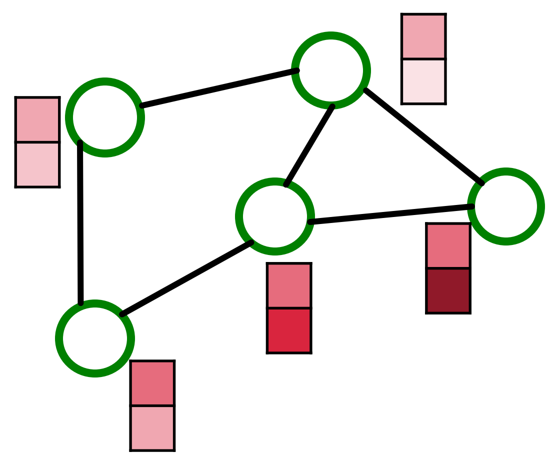
Graph Transformer Layer

J. You, R. Ying, J. Leskovec. Design Space of Graph Neural Networks, NeurIPS 2020

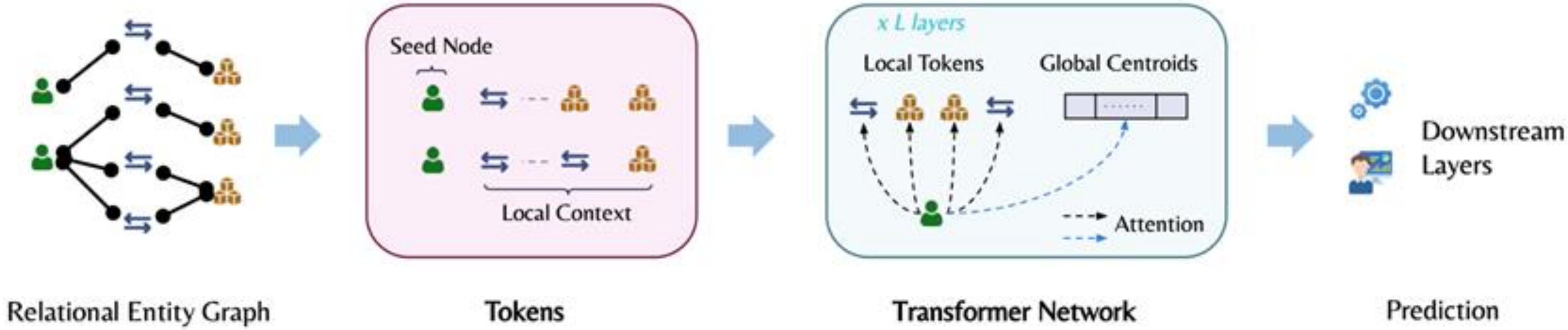
Processing Graphs with Transformers

- A graph Transformer must take the following inputs:
  - **(1) Node features?**
  - **(2) Adjacency information?**
  - ~~**(3) Edge features?**~~
  - **(3) Node types**
  - **(4) Timestamp**

- Key components of Transformer
  - **(1) tokenizing**
  - **(2) positional encoding**
  - **(3) self-attention**



Relational Graph Transformer



In **Relational Graph Transformer**

- We use multiple elements to represent the graph structure
  - Node feature, <multiple information elements>

structure

Limitations of existing PEs for RDL

- Designed for homogeneous graphs
- No temporal dynamics
- Computational overhead
- Graph changes over time

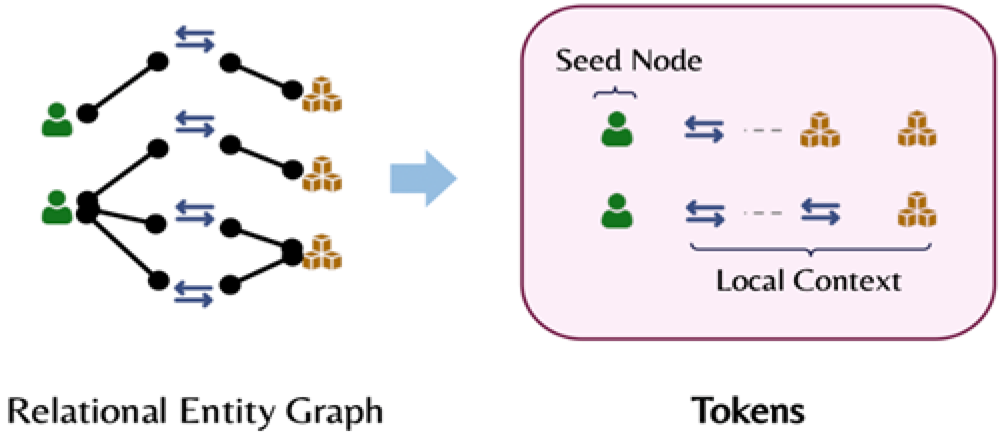
Unique properties of Relational Graphs

- The relational entity graph is:
  - Large-scale
  - Heterogeneous
  - Temporal
- Add a categorical encoding to capture the heterogeneity
- Add a time encoding to capture the temporal structure

How to encode structure

- **Challenge:** The relational entity graph is
  - Large-scale
- **Solution:** We need to sample a subgraph around the seed node

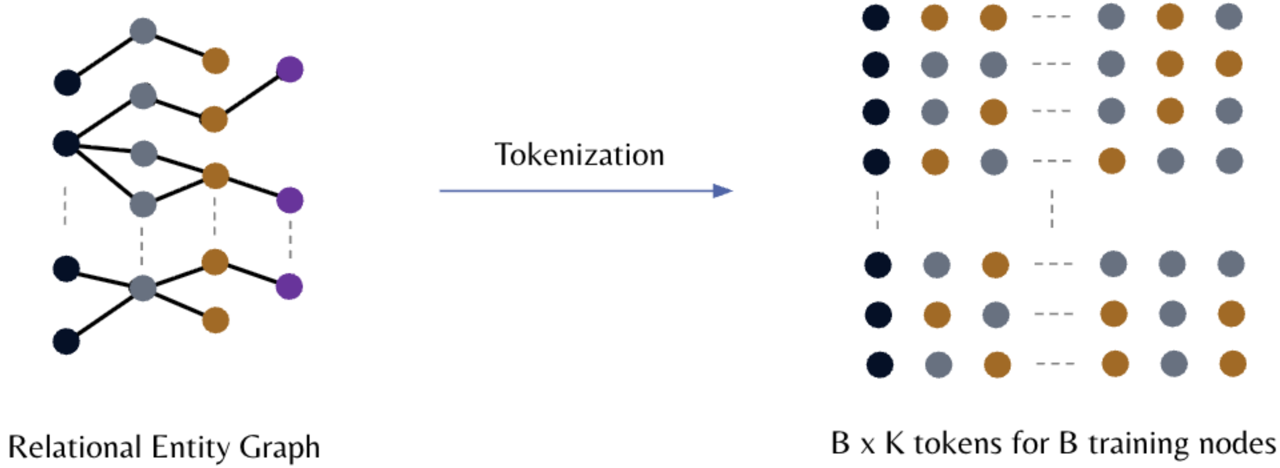
Relational Graph Transformer



Relational Entity Graphs

- **Training (seed) nodes** which correspond to node-types with respect to which tasks are defined.
- We are going to sample a context (subgraph) around each seed node.

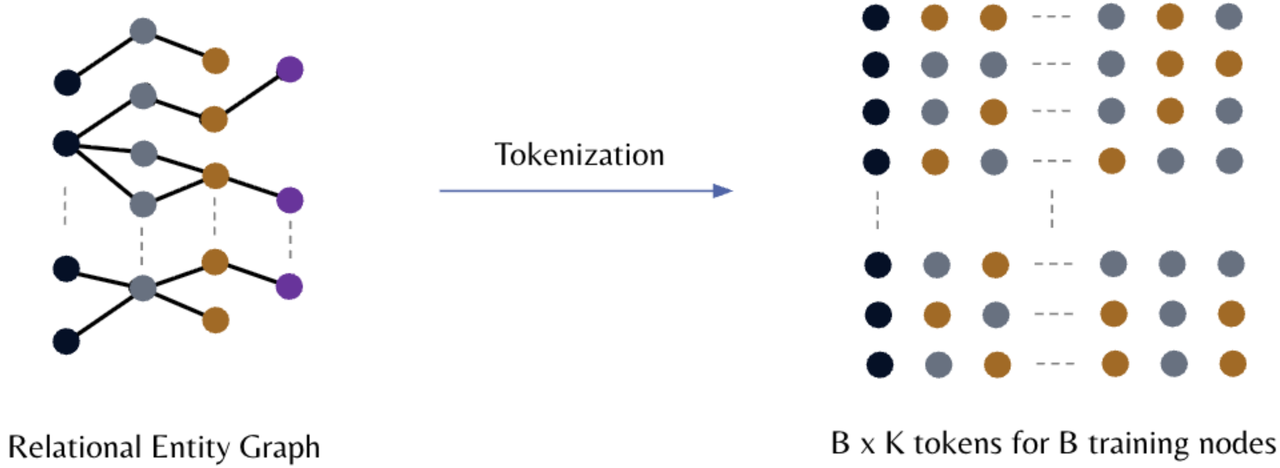
Relational Graph Transformer



Token Preparation (Sampling Stage)

- For each training node (•), a fixed set of ***K*** tokens (••) from local neighborhood (e.g., up to 2 hops) is selected through a temporal-aware sampling.
- We encode structure with a **GNN positional encoding** for each node in the context + **hop distance from seed node**.

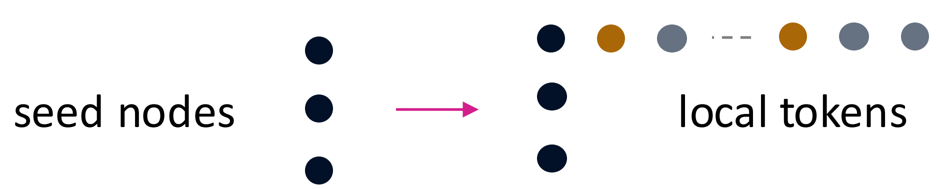
Relational Graph Transformer



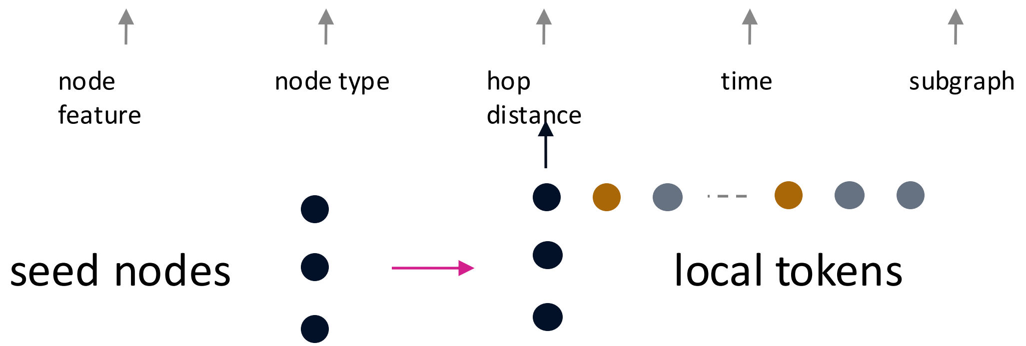
Token Preparation (Sampling Stage)

- For each training node (•), a fixed set of ***K*** tokens (••) from local neighborhood (e.g., up to 2 hops) is selected through a temporal-aware sampling.
- Each node in ***K*** set is represented by a ***5***-tuple: (node feature, node type, hop distance, time, GNN PE)

Relational Graph Transformer



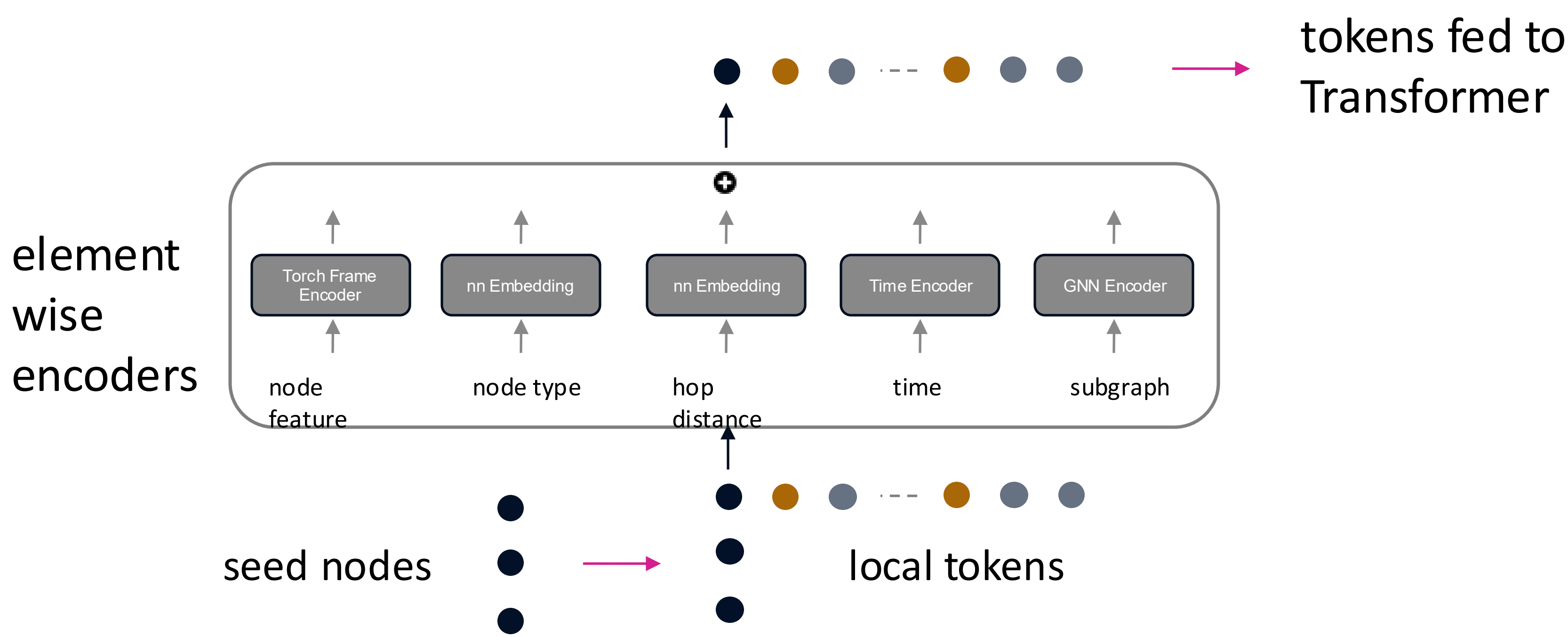
Relational Graph Transformer



Token Feature Encoders

- Each node in ***K*** set is represented by a ***5***-tuple:
  (node feature, node type, hop distance, time, GNN PE)

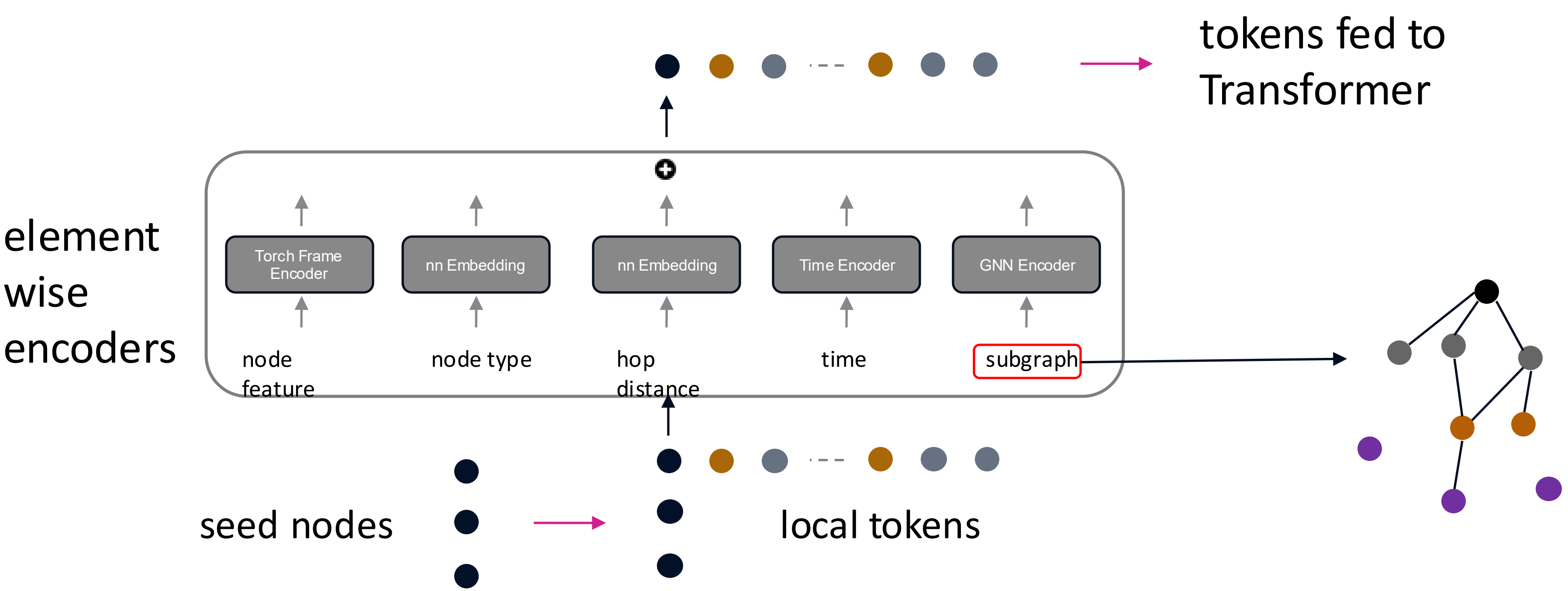
Relational Graph Transformer



Token Feature Encoders

- Each node in ***K*** set is represented by a ***5***-tuple: (node feature, node type, hop distance, time, GNN PE)
- Each of these elements is passed to a **feature encoder** and the combination becomes the token features for Transformer Network

Relational Graph Transformer



Token Feature Encoders

- Each node in ***K*** set is represented by a ***5***-tuple: (node feature, node type, hop distance, time, GNN PE)
- Each of these elements is passed to a **feature encoder** and the combination becomes the token features for Transformer Network

Ablation studies

Takeaway:

- Contribution of the multi-element tokenization strategy to capture arbitrary structure in relational data

Table 2: Relative drop (%) in performance in RELGT after removing a model component. Negative scores suggest the component is critical in RELGT, and vice-versa. Full results in Table 7.

Dataset	Task	No Global Module	No GNN PE	No Node Type	No Hop Distance	No Relative Time
rel-avito	ad-ctr	−6.00	−1.14	−7.14	−3.43	−9.14
rel-avito	user-clicks	7.79	−15.19	4.96	5.72	8.32
rel-avito	user-visits	−0.32	−2.35	−0.08	0.42	−0.72
rel-event	user-ignore	−1.28	0.15	−0.09	0.69	−0.06
rel-trial	study-outcome	−2.08	−1.66	3.80	−0.37	2.54
rel-trial	site-success	−19.08	−9.23	−2.94	−21.56	−0.77
rel-amazon	user-churn	−0.64	−0.78	0.16	0.06	−2.20
rel-hm	item-sales	−9.33	−17.35	−12.69	0.93	−77.24
	Average	−3.87	−5.95	−1.75	−2.19	−9.91

Results

Takeaway:

- Relational Graph Transformers improve over GNNs

Table 1: Test set results on the entity regression and classification tasks in RelBench. Best values are in **bold**. RDL: HeteroGNN baseline [40], HGT: Heterogeneous GT [20], PE: Laplacian Positional Encodings [9]. Relative gains are expressed as percentage improvement over RDL baseline.

(a) MAE for entity regression. Lower is better

Dataset	Task	RDL	HGT	HGT +PE	RelGT (ours)	% Rel. Gain
rel-f1	driver-position	4.022	4.1598	4.2358	**3.9170**	2.61
rel-avito	ad-ctr	0.041	0.0441	0.0494	**0.0345**	15.85
rel-event	user-attendance	0.258	0.2635	0.2562	**0.2502**	2.79
rel-trial	study-adverse	44.473	43.3253	**42.4622**	43.9923	1.08
	site-success	0.400	0.4374	0.4431	**0.3263**	18.43
rel-amazon	user-ltv	14.313	15.3804	15.9296	**14.2665**	0.32
	item-ltv	50.053	56.1384	55.6211	**48.9222**	2.26
rel-stack	post-votes	**0.065**	0.0679	0.0680	**0.0654**	-0.62
rel-hm	item-sales	0.056	0.0655	0.0641	**0.0536**	4.29

(b) AUC for entity classification. Higher is better.

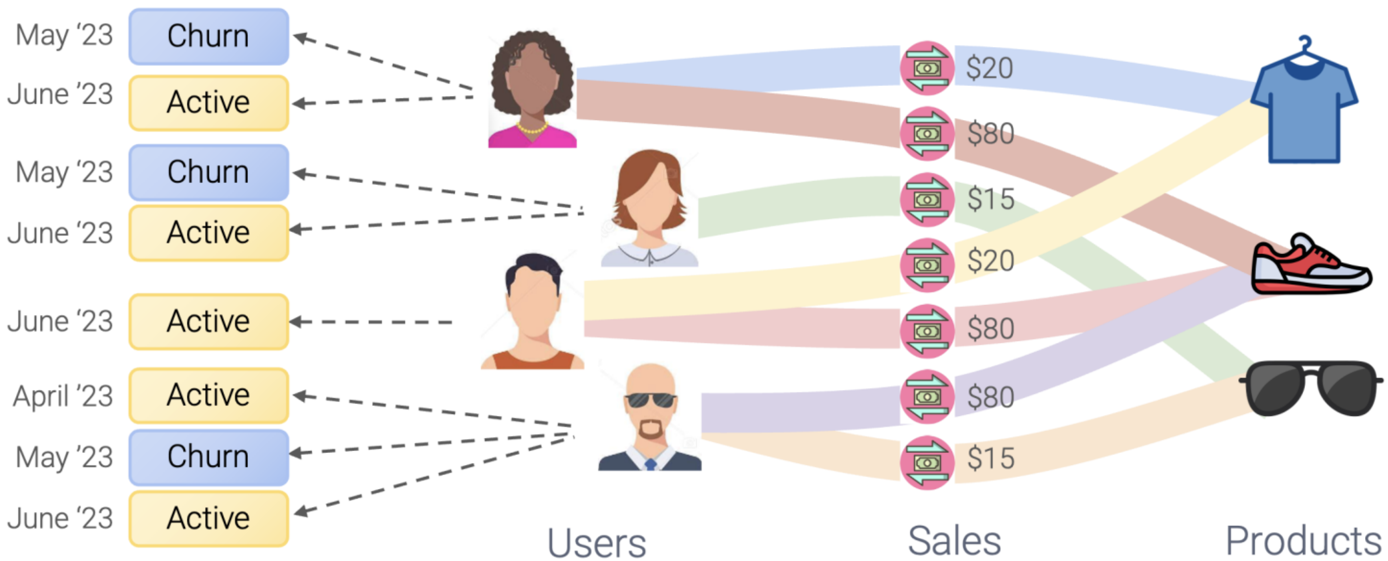
Dataset	Task	RDL	HGT	HGT +PE	RelGT (ours)	% Rel. Gain
rel-f1	driver-dnf	0.7262	0.7142	0.7109	**0.7587**	4.48
	driver-top3	0.7554	0.6389	0.8340	**0.8352**	10.56
rel-avito	user-clicks	0.6590	0.6584	0.6387	**0.6830**	3.64
	user-visits	0.6620	0.6426	0.6507	**0.6678**	0.88
rel-event	user-repeat	**0.7689**	0.6717	0.6590	0.7609	-1.04
	user-ignore	0.8162	**0.8348**	0.8161	0.8157	-0.06
rel-trial	study-outcome	0.6860	0.5679	0.5691	**0.6861**	0.01
rel-amazon	user-churn	**0.7042**	0.6608	0.6589	0.7039	-0.04
	item-churn	**0.8281**	0.7824	0.7840	0.8255	-0.31
rel-stack	user-engagement	0.9021	0.8898	0.8852	**0.9053**	0.35
	user-badge	**0.8986**	0.8652	0.8518	0.8632	-3.94
rel-hm	user-churn	**0.6988**	0.6773	0.6491	0.6927	-0.87

Stanford CS224W: Relational Foundation Model

CS224W: Machine Learning with Graphs
Jure Leskovec, Stanford University
http://cs224w.stanford.edu

Limitations of previous models

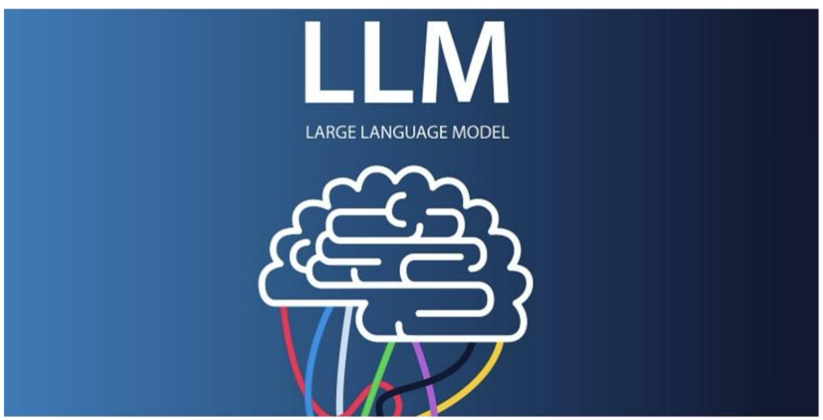
- They are schema- and task- specific

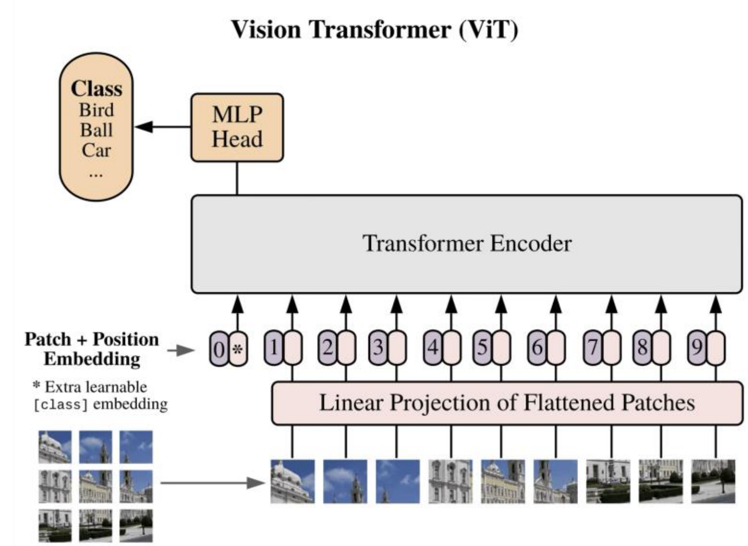


- The do not transfer knowledge from one database to another or one task to another
- Do not generalize to new databases or tasks

Can we do better?

- Goal: Bring the success of foundation models (LLMs, ViT) to relational data.





Relational Foundation Model

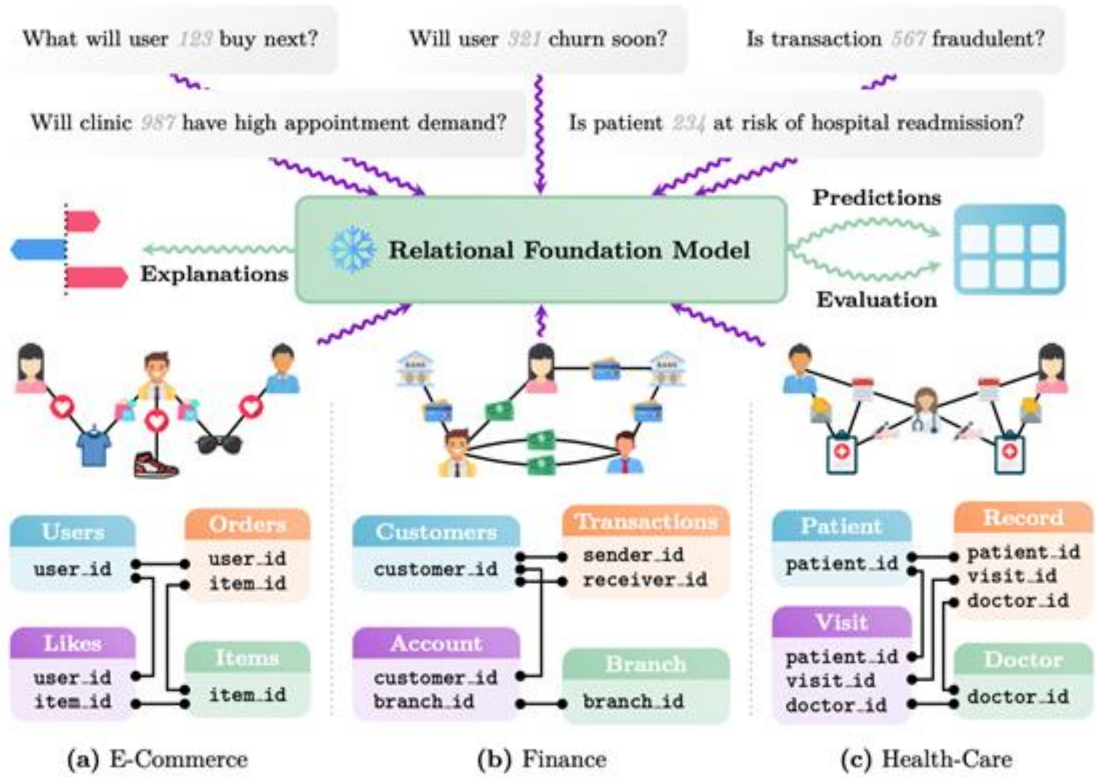
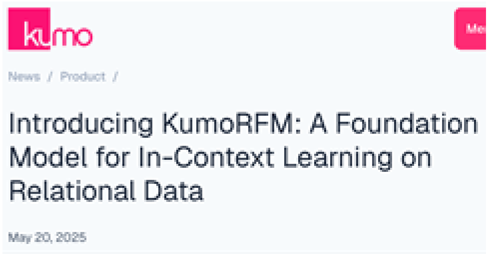
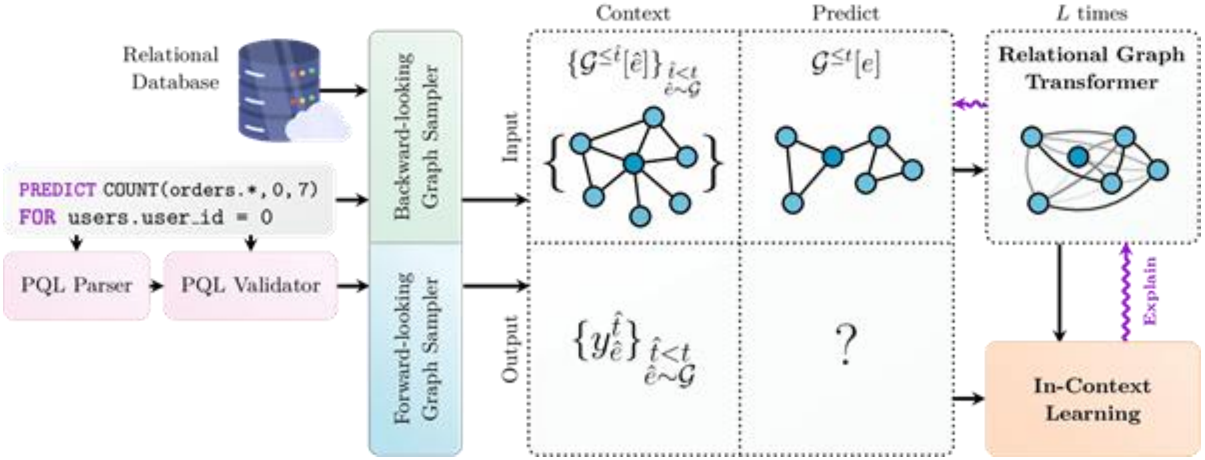


Figure 1: **Key capabilities of Relational Foundation Models.** RFMs can be applied to new/unseen databases and schemas with highly varying structural characteristics, as found in **(a)** e-commerce, **(b)** finance, or **(c)** health care. Secondly, they can be applied to any predictive task type, ranging from one-off assessments (*e.g.* entity-level fraud prediction) to temporal predictive queries (*e.g.* temporal recommendation prediction). Thirdly, they generalize to new predictive tasks and give accurate predictions without any task-specific model tuning. Finally, not only do RFMs support prediction outputs, but they also offer insights into the reasoning processes via explanations, and build trust through extensive quantitative evaluation mechanisms.

RelGT as the backbone for RFM





Relational Graph Transformer is the backbone for

first Relational Foundation Model

with zero-shot capabilities powered by ICL

Relational Foundation Model

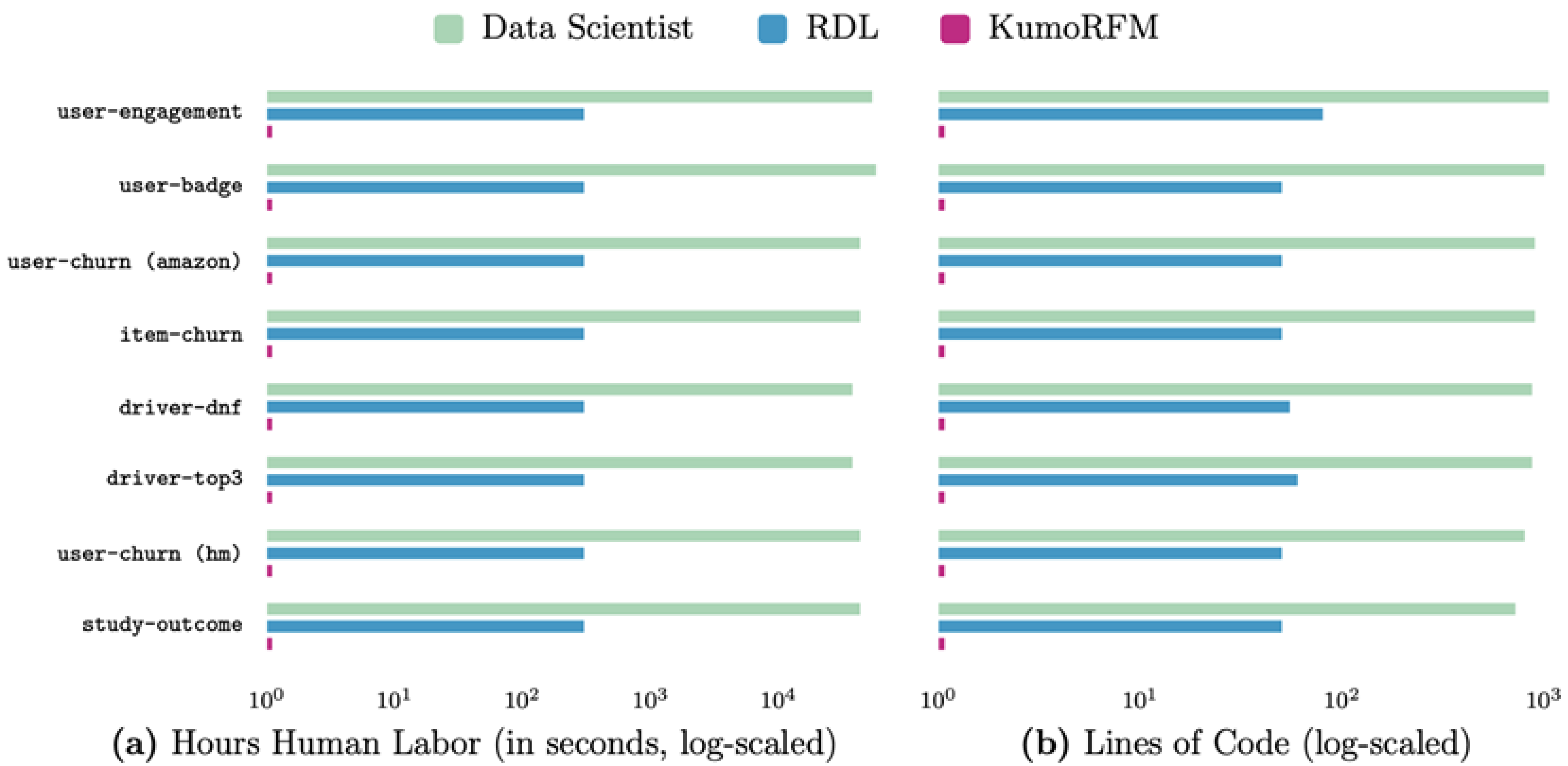


Figure 4: **Time to first prediction in (a) hours of human labor and (b) lines of code (LoC), illustrated on node-level classification tasks.** KumoRFM is orders of magnitude faster ($\approx$ 1 second *vs.* $\approx$ 30 minutes *vs.* $\approx$ 12.3 hours) and requires zero-code to get to accurate predictions (1 LoC *vs.* $\approx$ 56 LoC *vs.* $\approx$ 878 LoC) compared to Data Scientist and RDL baselines, respectively.